

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification 6 :</b> <b>A01N 63/00</b>	<b>A3</b>	<b>(11) International Publication Number:</b> <b>WO 95/01098</b> <b>(43) International Publication Date:</b> 12 January 1995 (12.01.95)
<b>(21) International Application Number:</b> PCT/US94/07252 <b>(22) International Filing Date:</b> 24 June 1994 (24.06.94) <b>(30) Priority Data:</b> 08/083,948 28 June 1993 (28.06.93) US <b>(71) Applicant:</b> MONSANTO COMPANY [US/US]; 800 North Lindbergh Boulevard, St. Louis, MO 63167 (US). <b>(72) Inventors:</b> CORBIN, David, Richard; 14453 Britannia Drive, Chesterfield, MO 63017 (US). GREENPLATE, John, Thomas; 747 Sherwick Terrace, Manchester, MO 63021 (US). JENNINGS, Michael, Girard; Apartment 307, 14550 Rialto Drive, Chesterfield, MO 63017 (US). PURCELL, John, Patrick; 615 Charbray Drive, Ballwin, MO 63011 (US). SAMMONS, Robert, Douglas; P.O. Box 69, New Melle, MO 63365 (US). <b>(74) Agent:</b> BOLDING, James, Clifton; Monsanto Company, 800 North Lindbergh Boulevard, St. Louis, MO 63167 (US).		<b>(81) Designated States:</b> AU, BB, BG, BR, BY, CA, CN, CZ, FI, GE, HU, JP, KE, KG, KR, KZ, LK, LV, MD, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>  <b>(88) Date of publication of the international search report:</b> 16 February 1995 (16.02.95)
<b>(54) Title:</b> METHOD OF CONTROLLING INSECTS  <b>(57) Abstract</b>  3-Hydroxysteroid oxidase controls insects. Genes encoding for this enzyme may be cloned into vectors for transformation of plant-colonizing microorganisms or plants, thereby providing a method of controlling insect infestation.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

# INTERNATIONAL SEARCH REPORT

International Application No.

PCT/85 94/07252

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 5 A01N63/00

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
IPC 5 A01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol.196, no.3, 15 November 1993, NEW YORK pages 1406 - 1403 J.P. PURCELL ET AL. 'Cholesterol oxidase: a potent insecticidal protein active against boll weevil larvae.' see the whole article -----	1-9

☐ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

\* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*Z\* document member of the same patent family

Date of the actual completion of the international search

5 December 1994

Date of mailing of the international search report

21.12.94

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+ 31-70) 340-3016

Authorized officer

DECORTE, D

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification<sup>6</sup> : A01N 63/00</p>	<p>A2</p>	<p>(11) International Publication Number: WO 95/01098 (43) International Publication Date: 12 January 1995 (12.01.95)</p>
<p>(21) International Application Number: PCT/US94/07252 (22) International Filing Date: 24 June 1994 (24.06.94) (30) Priority Data: 08/083,948 28 June 1993 (28.06.93) US (71) Applicant: MONSANTO COMPANY [US/US]; 800 North Lindbergh Boulevard, St. Louis, MO 63167 (US). (72) Inventors: CORBIN, David, Richard; 14453 Britannia Drive, Chesterfield, MO 63017 (US). GREENPLATE, John, Thomas; 747 Sherwick Terrace, Manchester, MO 63021 (US). JENNINGS, Michael, Girard; Apartment 307, 14550 Rialto Drive, Chesterfield, MO 63017 (US). PURCELL, John, Patrick; 615 Charbray Drive, Ballwin, MO 63011 (US). SAMMONS, Robert, Douglas; P.O. Box 69, New Melle, MO 63365 (US). (74) Agent: BOLDING, James, Clifton; Monsanto Company, 800 North Lindbergh Boulevard, St. Louis, MO 63167 (US).</p>		<p>(81) Designated States: AU, BB, BG, BR, BY, CA, CN, CZ, FI, GE, HU, JP, KE, KG, KR, KZ, LK, LV, MD, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p><b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i></p>
<p>(54) Title: METHOD OF CONTROLLING INSECTS</p> <p>(57) Abstract</p> <p>3-Hydroxysteroid oxidase controls insects. Genes encoding for this enzyme may be cloned into vectors for transformation of plant-colonizing microorganisms or plants, thereby providing a method of controlling insect infestation.</p>		



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## METHOD OF CONTROLLING INSECTS

### FIELD OF THE INVENTION

This invention relates to a method of controlling lepidopteran insects by use of an enzyme which may be applied directly to the plant or produced thereon by microorganisms or by genetically modifying the plant to produce the enzyme, and to genes, microorganisms, and plants useful in that method.

### BACKGROUND OF THE INVENTION

The use of natural products, including proteins, is a well known method of controlling many insect pests. For example, endotoxins of *Bacillus thuringiensis* (*B.t.*) are used to control both lepidopteran and coleopteran insect pests. Genes producing these endotoxins have been introduced into and expressed by various plants, including cotton, tobacco, and tomato. There are, however, several economically important insect pests that are not susceptible to *B.t.* endotoxins. There is also a need for additional proteins which control insects for which *B.t.* provides control in order to manage any development of resistance in the population.

It is therefore an object of the present invention to provide proteins capable of controlling lepidopteran insects and genes useful in producing such proteins. It is a further object of the present invention to provide genetic constructs for and methods of inserting such genetic material into microorganisms and plant cells. It is another object of the present invention to provide transformed microorganisms and plants containing such genetic material.

### SUMMARY OF THE INVENTION

It has been discovered that proteins that catalyze the oxidation of 3-hydroxysteroids, for example, cholesterol, will control lepidopteran insects. They cause mortality and stunting of larvae of lepidopteran insects. The enzymes may be applied directly to plants or introduced in other ways such as through the application of plant-colonizing microorganisms or by the plants themselves, which have been transformed to produce the enzymes.

3-Hydroxysteroid oxidases (E.C.1.1.3.6) are naturally produced by microorganisms such as *Streptomyces* sp., *Pseudomonas* sp., *Mycobacterium* sp., *Schizophyllum commune*, *Nocardia* sp., and *Rhodococcus* sp. [Smith et al., 1976, and Long et al., 1990.].

- 5 Preparations of enzymes from several different sources are available from Sigma Chemical Company, St. Louis, Missouri.

New *Streptomyces* genes that control the expression of 3-hydroxysteroid oxidase have been isolated and sequenced. These new genes or genes from other known producers of 3-hydroxysteroid oxidase  
10 may be inserted into a transformation vector cassette which is used to transform plant-colonizing microorganisms which when applied to plants express the genes producing a 3-hydroxysteroid oxidase, thereby providing control of insects. Alternatively, genes which function in plants and encode the subject enzymes may be inserted into  
15 transformation vector cassettes which may be incorporated into the genome of the plant, which then protects itself from attack by expressing the gene and producing a 3-hydroxysteroid oxidase. Additionally, the plant may also be transformed to co-express *B.t.* genes which express proteins for the control of other insects. Examples of  
20 plants transformed to express *B.t.* genes are disclosed in European Patent Publication No. 0 385 962 (Fischhoff et al.).

In accomplishing the foregoing, there is provided, in accordance with one aspect of the present invention, a method of controlling insect infestation of plants comprising providing a 3-hydroxysteroid oxidase for  
25 ingestion by the insect.

In accordance with another aspect of the present invention, there is provided a method of producing genetically transformed plants which express an amount of a 3-hydroxysteroid oxidase effective to control lepidioteran insects, comprising the steps of:

- 30 a) inserting into the genome of a plant cell a recombinant, double-stranded DNA molecule comprising
- (i) a promoter which functions in plant cells to cause the production of an RNA sequence;
  - (ii) a structural coding sequence that encodes for  
35 3-hydroxysteroid oxidase; and

- (iii) a 3' non-translated region which functions in said plant cells to cause the addition of polyadenylate nucleotides to the 3' end of the RNA sequence,

wherein said promoter is heterologous with respect to the structural coding sequence and wherein said promoter is operatively linked with said structural coding sequence, which is in turn operably linked with said non-translated region;

- b) obtaining transformed plant cells; and  
c) regenerating from the transformed plant cells genetically transformed plants which express an insecticidally effective amount of sterol oxidase.

There is also provided, in accordance with another aspect of the present invention, bacterial and transformed plant cells that contain DNA comprised of the above-mentioned elements (i), (ii), and (iii).

As used herein, the term "controlling insect infestation" means reducing the number of insects which cause reduced yield, either through mortality, retardation of larval development (stunting), or reduced reproductive efficiency.

As used herein, the term "structural coding sequence" means a DNA sequence which encodes for a polypeptide, which may be made by a cell following transcription of the DNA to mRNA, followed by translation to the desired polypeptide.

#### DETAILED DESCRIPTION OF THE INVENTION

3-Hydroxysteroid oxidases catalyze the oxidation of the 3-hydroxy group of 3-hydroxysteroids to produce ketosteroids and hydrogen peroxide. They are capable of catalyzing the oxidation of various 3-hydroxysteroids, such as, for example, cholesterol. Most of the previously known 3-hydroxysteroid oxidases are called "cholesterol oxidases" (enzymatically catalogued as E.C. #1.1.3.6) but cholesterol is only one of the 3-hydroxysteroid substrates. The use of all 3-hydroxysteroid oxidases and the genes encoding such proteins, for the purpose of controlling insects, are within the scope of the present invention.

3-Hydroxysteroid oxidases are commercially available for use as reagents for serum cholesterol assays. For example, Sigma Chemical

Company, St. Louis, Missouri, offers three such 3-hydroxysteroid oxidases (denominated as cholesterol oxidases), one from a *Streptomyces* sp., one from a *Pseudomonas fluorescens*, and one from a *Brevibacterium*. Two other sources of 3-hydroxysteroid oxidase, two streptomycetes  
5 denominated A19241 and A19249, each of which produce a 3-hydroxysteroid oxidase, have been isolated. The organisms were collected in Madagascar. When these organisms were cultured according to usual methods the culture filtrates were found to affect insect larvae as described below.

10 A seed culture of A19249 was started in 55 mL sterile Tryptone-Yeast Extract broth, pH 6.8, in a 250 mL shake flask. The seed was agitated at 250 rpm on a rotary shaker for 3 days at 30 °C. A New Brunswick Bioflo II Bioreactor with a 2 L working volume was filled with "medium 202" ( $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  2 g/L,  $\text{KH}_2\text{PO}_4$  0.5 g/L, NaCl 0.5 g/L,  
15  $\text{CaCO}_3$  1 g/L,  $\text{ZnSO}_4 \cdot \text{H}_2\text{O}$  (1 mg/mL stock) 5 mL/L, 100 mM FeEDTA 0.5 mL/L, Soluble Starch 5 g/L, Dextrose 2.5 g/L, Malt Extract 2.5 g/L, Soytone 5 g/L). The pH was adjusted to 6.5 with 2.5 N NaOH or 1 N HCl, and 1 mL/L of P2000, an antifoam agent was added. The bioreactor was sealed and autoclaved for 25 min at 250 °C. The seed, at  
20 3 days growth, was used to inoculate the fermentor at 2% or 40 mL. The fermentation took place at 30 °C with an airflow of 1 L/min and agitation running at 500 rpm. The fermentation was harvested after 40 h.

Each of these enzymes has demonstrated control of insects as  
25 shown below. The *P. fluorescens* 3-hydroxysteroid oxidase is immunologically distinct from the *Streptomyces* enzymes, but it also controls insects.

Other organisms producing 3-hydroxysteroid oxidases of the present invention may be identified by assaying culture filtrates or  
30 individual proteins for 3-hydroxysteroid oxidase activity using a spectrophotometric assay, described below, which measures hydrogen peroxide production in the presence of a 3-hydroxysteroid, for example, cholesterol [Gallo, 1981].

## BIOEFFICACY ASSAYS

### Larval Bioassay

Lepidopteran larvae were tested on artificial agar-based diet (Marrone et al., 1985) treated with the indicated amount of the A19249 3-hydroxysteroid oxidase (cholesterol oxidase) for six days. The results are shown in Table 1. Percent stunting is defined as

$$\frac{\text{larval weights (Hepes control)} - \text{larval weights (cholesterol oxidase)}}{\text{larval weights (Hepes control)}} \times 100\%$$

10      larval weights (Hepes control)

An extended test was performed with tobacco budworm larvae to test the effect of the stunting noted in the six-day test. Tobacco budworm eggs were added to artificial diet (as described above) containing either buffer or 100 ppm A19249 3-hydroxysteroid oxidase (cholesterol oxidase). After seven days, some mortality as compared to the controls was noted. Surviving larvae were moved to fresh diet (control or treated, as appropriate) every seven days. Percent mortality (corrected for control mortality) is reported for the 7, 10, and 14 day periods in Table 2. The corrected number of larvae was 23. By the 27th day only two larvae were alive; both were very stunted and never pupated. 92% of the control larvae survived and all survivors pupated by day 27. This experiment demonstrates that 3-hydroxysteroid oxidase essentially arrests the development of tobacco budworm larvae.

25

Table 1

<u>Insect</u>	<u>Stage</u>	<u>Dose-<math>\mu</math>g/mL</u>	<u>Stunting</u>
tobacco budworm	egg/lv	30	0
	lv	100	86%
corn earworm	lv	50	0
	lv	100	35%
fall army worm	lv	30	0
tobacco hornworm	lv	30	0
	lv	100	30%
pink bollworm	lv	50	0
	lv	100	30%

35

6

European cornborer	lv	50	0
	lv	100	46%
beet armyworm	lv	100	76%
black cutworm	lv	100	68%

5

Table 2

<u>Interval (days)</u>	<u>Percent Mortality</u>
7	20
10	61
10 14	80

### MODE OF ACTION STUDIES

The following studies show that 3-hydroxysteroid oxidase has a direct effect on the insect itself and that the activity demonstrated in the experiments described above cannot be attributed to the enzymes effect on the insect's diet, for example by sterol depletion. Boll weevil and lepidopteran larvae are most susceptible to the enzyme. It is believed that this specificity is due to the effect of 3-hydroxysteroid oxidase on the midgut of the insect as explained in more detail below.

Other insects with similar midgut physiologies may also be controlled by 3-hydroxysteroid oxidase. In addition, 3-hydroxysteroid oxidases other than those tested and reported herein may control a different spectrum of insects with different midgut physiologies.

#### 25 Cotton Seed Diet Assay

Treatment diet was made by mixing 30 g of Pharmamedia™ (Traders Protein) cottonseed flour into 170 mL of a 1.6% agar solution at 50 °C, containing 0.13% propionic acid, 0.014% phosphoric acid, and 30 mg each of streptomycin sulfate and chlor-tetracycline. Before mixing, 10% KOH was used to adjust the pH to 6.2. Pharmamedia™ is a flour made up of cottonseed embryos. The diet was incubated in a water bath at 40 °C. Dilutions of the Sigma *Streptomyces* 3-hydroxysteroid oxidase (cholesterol oxidase) were incorporated into the diets. Tobacco budworm larvae are 68% stunted when exposed to 3-hydroxysteroid oxidase (100 ppm) in cottonseed diet.

### Cotton Callus Diet Assay

Cotton callus assays were conducted using Coker 312 cotton callus and a 96 well insect diet tray. Each well contained 0.5 ml of gelled  
5 2% agar (containing 0.13% propionic acid and 0.014% phosphoric acid) that was covered with a filter paper disc. A piece of callus (approximately 150 mg) was placed in each well. For each replicate, 25  $\mu$ l of a stock solution of 3-hydroxysteroid oxidase (1.25 mg/ml stock) or 25  $\mu$ l of 25 mM Hepes buffer (pH 7.5) was pipetted onto each piece of  
10 callus. Tobacco budworm eggs (4-8 eggs/well, 0-12 hr from hatching) in 0.15% agar were pipetted onto the filter disc next to the callus. Individual larvae were transferred to freshly treated callus samples every three days. Larvae were weighed after 10 days.

Stunting of larvae was observed in the treated wells, although  
15 reduced as compared to the artificial diet or the cottonseed diet. This reduced level of stunting with cotton callus is probably due to the pipetted enzyme solution not completely covering the callus tissue and the loss of some enzyme onto the filter paper.

### 20 Pre-ingestion Effect on Diet

There are no lethal or stunting effects from feeding larvae a diet sample that was incubated for one week with a 3-hydroxysteroid oxidase and then heated to 80 °C to denature the enzyme prior to using it in the  
above assay. This further demonstrates that the mode of action of  
25 3-hydroxysteroid oxidase is not dependent on sterol depletion of the nutrient source but that the enzyme is directly active upon the insect.

### Microscopy Studies

Tobacco budworm larvae were reared from eggs on artificial diet  
30 containing 100  $\mu$ g/ml of 3-hydroxysteroid oxidase or Hepes buffer control. After ten days, midgut epithelia were dissected and processed for microscopy as described by Purcell et al., 1993. The midguts from treated larvae exhibited a variety of histological and ultrastructural alterations as compared with Hepes buffer-treated controls. 3-Hydroxy-  
35 steroid oxidase induced a lateral constriction and apical-basal elongation



of the cells in the midgut epithelium, resulting in a widening of the inter-cellular spaces and the formation of an irregular and convoluted luminal surface. The cellular attenuation was also characterized by localized apical cytoplasmic blebbing and microvillar denudation. This denudation occasionally generated detached cytoplasmic fragments in the lumen. Localized cytolysis was observed, but the epithelium as a whole remained intact. These observations suggest that exposure to 3-hydroxysteroid oxidase compromises the integrity of the apical plasma membrane which results in a secondary generalized cellular hypersensitivity to osmotic swelling and lysis.

Cholesterol is required for the integrity and normal function of virtually all cellular membranes. A reduction in the availability of cholesterol for incorporation into and maintenance of cellular membranes due to metabolism by 3-hydroxysteroid oxidase could therefore be one possible mechanism for toxicity of this agent. However, the observations described above show the most dramatic structural alterations are localized to the tips of the midgut cells in contact with the 3-hydroxysteroid oxidase in the gut lumen. This suggests that 3-hydroxysteroid oxidase may act directly to alter cholesterol already incorporated into the membrane. In addition, diet treatment studies described above demonstrated that the diet was not altered prior to ingestion in a manner which compromises tobacco budworm growth, again suggesting a direct effect of the enzyme on the insect.

#### Mode of Action Theory

While not being bound by this theory, it is believed that the 3-hydroxysteroid oxidase enzyme stunts the growth of lepidopteran larvae by some action in the gut after ingestion. The bioassay and morphological data show that there was clearly a pathological condition resulting from ingestion of diet containing 3-hydroxysteroid oxidase.

### ENZYME IDENTIFICATION

The active proteins from the Madagascar *Streptomyces* microorganisms were isolated, purified, partially sequenced, and identified as 3-hydroxysteroid oxidases.

### Protein Isolation

Each culture filtrate was purified by first sizing on YM10 membranes (Amicon) to a [ $>10$  kDa] fraction, followed by multiple chromatography runs on an FPLC Mono Q HR10/10 (Pharmacia LKB, Piscataway, NJ) column. For chromatography on the Mono Q column, the samples were loaded on the column in 25 mM Hepes pH 7.5 and eluted with a gradient to 1.0 M KCl in 25 mM Hepes pH 7.5. Fractions were collected and aliquots of each were filtered through 0.2  $\mu$  Acrodisc syringe tip filters. Each was tested in an insect assay. Aliquots of insecticidally active fractions were electrophoresed on SDS-PAGE [Laemmli, 1970] using a Daiichi Double Gel Device and 10-20% mini-gel. Proteins were visualized by silver staining using Daiichi silver stain reagent kit. The active enzymes of the present invention, isolated from the novel microorganisms, were found to be a 52.5 kDa protein.

### Amino Acid Sequences

An SDS-PAGE gel of the protein produced by *Streptomyces* A19241, isolated as above, was blotted onto PVDF paper (Immobilon, Millipore) using the protocol of Matsudaira [Matsudaira, 1987]. The N-terminus was sequenced using automated Edman degradation chemistry. A gas phase sequencer (Applied Biosystems, Inc.) was used for the degradation using the standard sequencer cycle. The respective PTH-aa derivatives were identified by reverse phase HPLC analysis in an on-line fashion employing a PTH analyzer (Applied Biosystems, Inc.) fitted with a Brownlee 2.1 mm i.d. PTH-C18 column. For internal sequences, digestions were carried out on purified 3-hydroxysteroid oxidase from A19249 using trypsin (TPCK-treated, from Worthington Biochemicals Corp., Freehold, NJ). Fragments were then purified by reverse phase HPLC and sequenced in an N-terminal fashion.

The resulting partial sequences were compared to known proteins and a strong (71%) homology was found with the reported fourteen amino acid sequence at the N-terminus of a 3-hydroxysteroid oxidase isolated from a *Streptomyces* species [Ishizaki et al., 1989]. The reported enzyme has an  $M_r$  of 54.9 kDa which agrees well with the  $M_r$  of 52.5 kDa of the isolated enzyme.

Six internal fragments of the purified enzyme from A19249, also having homology to six regions of the reported enzyme, were sequenced. The fragments had 95, 76, 64, 58, 89, and 100 percent sequence identities.

5

#### Amino Acid Composition Determination and Comparison

The amino acid composition of the 3-hydroxysteroid oxidase produced by A19249 was determined and compared with the composition of the reported *Streptomyces* enzyme. The samples were subjected to acid hydrolysis (6N HCl, vapor phase hydrolysis using a Water's Picotag workstation, 24 hr, 110 °C.). All analyses were performed after post-column derivation of the hydrolysates using ninhydrin [Moore et al., 1963]. A Beckman Model 6300 Auto analyzer was employed for the actual determinations. The S delta n/N statistic is used to compare two compositions in order to make a prediction about their relatedness. The formula for the statistic is:

$$1/2 \sum (n_{Ai} - n_{Bi})^2/N$$

where A is one composition, B is the other composition, i is each amino acid, n is the number of each amino acid, N is the total number of amino acids in the protein. If S delta n/N is <0.42, then there is a greater than 95% chance that the proteins are related. The smaller the value, the more closely the determined compositions match.

The S delta n/N statistic for the A19249 protein compared to the reported enzyme is 0.36, indicating that the two are highly related.

25

#### 3-Hydroxysteroid Oxidase Assay

The identity of the enzyme was confirmed by testing its ability to oxidize a 3-hydroxysteroid, specifically cholesterol. The enzyme is added to a reagent mixture comprising horseradish peroxidase (20 U/mL), phenol (14 mM), 4-amino antipyrine (0.82 mM), Triton® X-100 (0.05%) and phosphate buffer (100 mM, pH 7). The sterol in isopropanol is then added and the absorbance at 500 nm monitored. One unit of activity is defined as the amount of enzyme required to oxidize 1 µmole of sterol per minute at 20 °C.

35

The activity levels of the enzymes are reported in Table 12 for 3-

hydroxysteroids representative of various classes of 3-hydroxysteroids.

The enzyme sources are as follows:

1 = A19249

2 = A19241

5 3 = Sigma *Streptomyces*

4 = Sigma *Pseudomonas*

Table 3

		<u>Relative Rate for Enzymes</u>			
<u>Sterol</u>		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>
10	cholesterol	100	100	100	100
	dihydrocholesterol	56	56	59	69
	dehydrocholesterol	13	12	7	47
	lathosterol	28	34	27	71
	stigmasterol	22	28	11	21
15	sitosterol	88	65	49	50
	campesterol*	65	64	45	49
	fucosterol	22	20	12	68
	lanosterol	<1	<1	<1	1
	ecdysone	<1	<1	<1	<1
20	20-OH ecdysone	<1	<1	<1	<1

\*65/35 mixture of campesterol and dihydrobrassicasterol

#### Immunological Comparison of Enzymes

The Sigma *Streptomyces* enzyme is immunologically related to the  
 25 3-hydroxysteroid oxidases produced by the isolates of the present  
 invention, numbers A19241 and 19249, as demonstrated by Western  
 blotting [Burnette et al., 1981] using polyclonal antisera generated  
 against the Sigma *Streptomyces* enzyme. The antisera recognized both  
 enzymes produced by the isolates. The 3-hydroxysteroid oxidase from *P.*  
 30 *fluorescens* was not recognized by the antisera. This demonstrates that  
 immunologically distinct 3-hydroxysteroid oxidases are toxic to insects.

#### **GENETIC IDENTIFICATION**

The 3-hydroxysteroid oxidase gene was isolated from one of the  
 35 *Streptomyces* microorganisms isolated in Madagascar and its sequence

determined.

### Cloning of the 3-Hydroxysteroid Oxidase Gene from A19249

As discussed above, peptide sequences of purified 3-hydroxysteroid  
5 oxidase from A19249 were obtained for four regions of the protein. These  
peptide sequences were compared to a database of known protein  
sequences, and this comparison revealed that the A19249 protein showed  
a high degree of homology to a known 3-hydroxysteroid oxidase from  
*Streptomyces* [Ishizaki]. Comparing the A19249 peptide sequences to this  
10 known protein sequence, these peptides were assigned to their likely  
positions in the A19249 protein sequence. The sequence derived from the  
intact 3-hydroxysteroid oxidase from A19249 corresponded to a region  
near the N-terminus of the secreted form of the enzyme from the pub-  
lished sequence. From this it was concluded that the A19249 N-terminal  
15 peptide sequence was also likely to correspond to a "mature" secreted  
form of the protein lacking its putative secretory signal sequence. This  
was later confirmed by the DNA sequence analysis of the A19249 gene  
(see below). Three peptides were used to construct hybridization probes  
for isolation of the A19249 3-hydroxysteroid oxidase gene. Peptide N2  
20 (SEQ ID NO:1) corresponded to N-terminal amino acids 29 - 43 of the  
known mature protein sequence (sequence without the putative signal  
peptide); peptide C1 (SEQ ID NO:2) to amino acids 434 - 449 of the  
mature protein sequence; and peptide C2 (SEQ ID NO:3) to amino acids  
464 - 475 of the mature protein sequence.

25 N2 (SEQ ID NO:1):

ValSerThrLeuMetLeuGluMetGlyGlnLeuTrpAsnGlnPro

C1 (SEQ ID NO:2):

AlaPheAlaAspAspPheCysTyrHisProLeuGlyGlyCysValLeu

C2 (SEQ ID NO:3): AsnLeuTyrValThrAspGlySerLeuIleProGly

30 Based on these peptide sequences, three long nondegenerate  
oligonucleotides, corresponding to 3-hydroxysteroid oxidase peptide  
sequences from A19249, were designed using *Streptomyces* preferred  
codons. The oligonucleotides N2 (SEQ ID NO:4), C1 (SEQ ID NO:5), and  
C2 (SEQ ID NO:6) correspond to the peptides N2, C1, and C2 described  
35 above.

N2 Probe (SEQ ID NO:4):

gtgtccaccctgatgctggagatgggcccagctgtggaaccagccc

C1 Probe (SEQ ID NO:5):

gccttcgccgacgacttctgctaccaccgctcggcggctgcgtctctg

5 C2 Probe (SEQ ID NO:6): aacctctacgtgaccgacggttcgctgatcccggt

Probes N2 (SEQ ID NO:4), C1 (SEQ ID NO:5), and C2 (SEQ ID NO:6) were all used as hybridization probes on Southern blots of A19249 genomic DNA. All three probes hybridized to the same 2.2 kb band in BamHI digested DNA, but N2 (SEQ ID NO:4) hybridized to a different  
10 fragment than C1 (SEQ ID NO:5) and C2 (SEQ ID NO:6) did in SalI and BglII digests. This indicated that SalI and BglII cut within the coding sequence of the 3-hydroxysteroid oxidase gene from A19249, which was confirmed by DNA sequence analysis.

The 3-hydroxysteroid oxidase gene from A19249 was isolated using  
15 the three synthetic oligonucleotides as hybridization probes on a library of DNA fragments of A19249 DNA in a lambda phage vector. A library was made in lambda EMBL3 using partial-digest MboI DNA fragments of A19249. These probes were used to screen approximately 72,000 lambda phage plaques from the primary library. Primary plaque screening was  
20 performed using N2 (SEQ ID NO:4) plus C2 (SEQ ID NO:6). A total of 12 recombinant plaques that hybridized to the N and C-terminal probes were picked and purified by a second round of hybridization screening with probes N2 (SEQ ID NO:4) and C2 (SEQ ID NO:6). Southern blot analysis revealed that, in five of six lambda clones analyzed, a 2.2 kb BamHI frag-  
25 ment hybridized to both the N and C-terminal probes. This result confirmed the earlier Southern hybridization analysis that indicated a 2.2 kb BamHI fragment contained the 3-hydroxysteroid oxidase gene. This 2.2 kb DNA fragment was cloned into plasmid vector pUC18 [Yanisch-Perron et al., 1985] in both orientations for further analysis. Restriction mapping  
30 showed that there were internal SalI and BglII sites as predicted by the Southern hybridization analysis. These sites are also conserved compared to the published 3-hydroxysteroid oxidase gene sequence. The BamHI fragment was further subcloned into four fragments for direct DNA sequencing.

### Sequence Analysis of the 3-Hydroxysteroid Oxidase Gene

A total of 1865 nucleotides of DNA sequence from the 2.2 kb BamHI fragment were determined by direct DNA sequence analysis of subclones of this fragment using the dideoxy chain termination method.

- 5 This sequence is identified as SEQ ID NO:7. This DNA sequence contains noncoding flanking regions at both the 3' and 5' ends. Analysis of this DNA sequence revealed a single long open reading frame that encodes a secretory signal peptide and the mature 3-hydroxysteroid oxidase protein of 43 and 504 amino acids, respectively. It is 84.37% identical to the
- 10 published 3-hydroxysteroid oxidase nucleotide sequence. The derived amino acid sequence is 81.685% identical to the published 3-hydroxysteroid oxidase sequence. It is identified as SEQ ID NO: 8. Examination of the A19249 DNA sequence and comparison to the N-terminal amino acid sequence of intact 3-hydroxysteroid oxidase from A19249 revealed that
- 15 the A19249 gene encoded a protein that includes a signal peptide sequence, which is apparently cleaved during secretion of the protein from the cells. Thus the N-terminus of the mature protein from A19249 begins with Ser-Gly-Gly-Thr-Phe, identified as SEQ ID NO:12.

20

### **GENETIC TRANSFORMATION**

- A 3-hydroxysteroid oxidase gene can be isolated from novel organisms or may be obtained from known sources, such as the *Rhodococcus* sp. described by Long et al., in WO 90 05,788. This gene may then be used to transform bacterial cells or plant cells to enable the production of 3-
- 25 hydroxysteroid oxidase and carry out methods of this invention. Examples of how this may be done with the gene of A19249 are given below.

### Mutagenesis of the A19249 Gene

- In order to incorporate the A19249 gene into vectors appropriate
- 30 for expression of the 3-hydroxysteroid oxidase in heterologous bacterial or plant hosts, it was necessary to introduce appropriate restriction sites near the ends of the gene. The goals of this mutagenesis were to create cassettes that included the protein coding sequence with minimal non-coding flanking sequences and to incorporate useful restriction sites to
- 35 mobilize these cassettes. Cassettes were designed that would allow mobili-

- zation of the intact coding sequence including the signal peptide or just the mature coding sequence. To incorporate these cassettes into appropriate bacterial or plant expression vectors, an NcoI restriction site was engineered at the N-terminus of the intact protein sequence or at the
- 5 N-terminus of the mature protein sequence. A BamHI site was engineered just after the termination codon of the intact coding sequence. Three mutagenesis primers were designed to create these cassettes, as shown below. Mutagenesis with primer Chossn (SEQ ID NO:9) substituted three amino acids (MAT) for valine and asparagine at the N-
- 10 terminus of the signal peptide of the intact protein and Chomnr (SEQ ID NO:10) added two amino acids (MA) at the N-terminus of the mature protein. This was necessary to allow incorporation of the NcoI restriction site. Mutagenesis with primer Cho3br (SEQ ID NO:11) incorporated a BamHI site at the 3' end of the coding sequence. Primers Chomnr and
- 15 Cho3br were used to direct formation of the antisense strand of DNA. Chossn (SEQ ID NO:9): CTCAGGAGCACCATGGCGACCGCACAC  
(NcoI site underlined)  
Chomnr (SEQ ID NO:10): GTGCCGCGGAGGCCATGGGGGCGGTGGC  
(NcoI site underlined)
- 20 Cho3br (SEQ ID NO:11):  
GCCCCGCCCCGTCGGATCCGTCAGGAACCCG (BamHI site underlined)  
The resulting modified sequences were identified as SEQ ID NO:13 encoding for the intact protein and SEQ ID NO:14 for the mature protein.

25 Expression of 3-Hydroxysteroid Oxidase in *E. coli*

- The NcoI-BamHI fragments containing either the intact protein coding sequence or the mature protein coding sequence were inserted into a vector designed for protein expression in *E. coli*, vector pKK233-2 (Pharmacia LKB, Piscataway, NJ). pKK233-2 contains the IPTG-
- 30 inducible trc promoter. The vector containing the intact (full length) protein coding sequence as modified (SEQ ID NO:13) is designated pMON20909. The vector containing the mature protein coding sequence as modified (SEQ ID NO:14) is designated pMON20907. *E. coli* XL1 Blue cells (Statagene, San Diego, CA) modified with pMON20909 expressed 3-
- 35 hydroxysteroid oxidase at higher levels of enzymatic activity than cells



modified with pMON20907. The protein was extracted and purified from 4 l of IPTG-induced *E. coli* containing pMON20909. The soluble fraction from sonicated bacterial lysate was concentrated and dialyzed, and then partially purified by Mono Q chromatography to yield 11 units of 3-hydroxy-steroid oxidase activity. Western blot analysis indicates that the signal sequence of the intact protein is cleaved in *E. coli*, but the exact site of cleavage was not determined. Analysis of the recovered protein showed a five-fold reduction in enzymatic activity relative to the A19249 protein, but the loss has not been explained by DNA sequencing which found no alterations that would explain loss of enzymatic activity in plant protoplasts or *E. coli*.

The recovered protein was tested against tobacco budworm and resulted in 88% stunting at a dose of 100 µg/ml.

#### 15 Expression of 3-Hydroxysteroid Oxidase in Plant Colonizing Bacteria

To control insects, it may be desirable to express 3-hydroxysteroid oxidase in plant colonizing bacteria, and then apply this bacteria to the plant. As the insect feeds on the plant, it ingests a toxic dose of 3-hydroxysteroid oxidase produced by the plant colonizers. Plant colonizers can be either those that inhabit the plant surface, such as *Pseudomonas* or *Agrobacterium* species, or endophytes that inhabit the plant vasculature such as *Clavibacter* species. For surface colonizers, the 3-hydroxysteroid oxidase gene may be inserted into a broad host range vector capable of replicating in these Gram-negative hosts. Examples of these such vectors are pKT231 of the IncQ incompatibility group [Bagdasarian et al., 1981] or pVK100 of the IncP group [Knauf, 1982]. For endophytes the 3-hydroxysteroid oxidase gene can be inserted into the chromosome by homologous recombination or by incorporation of the gene onto an appropriate transposon capable of chromosomal insertion in these endophytic bacteria.

30

#### Plant Gene Construction

The expression of a plant gene which exists in double-stranded DNA form involves transcription of messenger RNA (mRNA) from one strand of the DNA by RNA polymerase enzyme, and the subsequent processing of the mRNA primary transcript inside the nucleus. This processing

35

involves a 3' non-translated region which adds polyadenylate nucleotides to the 3' end of the RNA. Transcription of DNA into mRNA is regulated by a region of DNA usually referred to as the "promoter." The promoter region contains a sequence of bases that signals RNA polymerase to  
5 associate with the DNA and to initiate the transcription of mRNA using one of the DNA strands as a template to make a corresponding strand of RNA.

A number of promoters which are active in plant cells have been described in the literature. Such promoters may be obtained from plants  
10 or plant viruses and include, but are not limited to, the nopaline synthase (NOS) and octopine synthase (OCS) promoters (which are carried on tumor-inducing plasmids of *Agrobacterium tumefaciens*), the cauliflower mosaic virus (CaMV) 19S and 35S promoters, the light-inducible promoter from the small subunit of ribulose 1,5-bisphosphate carboxylase  
15 (ssRUBISCO, a very abundant plant polypeptide), and the Figwort Mosaic Virus (FMV) 35S promoter. All of these promoters have been used to create various types of DNA constructs which have been expressed in plants (see e.g., PCT publication WO 84/02913).

The particular promoter selected should be capable of causing  
20 sufficient expression of the enzyme coding sequence to result in the production of an effective amount of 3-hydroxysteroid oxidase. A preferred promoter is a constitutive promoter such as FMV35S. It has been observed to provide more uniform expression of heterologous genes in the flowering portions of plants. Use of such a promoter with the 3-hydroxysteroid  
25 oxidase gene may provide greater protection of cotton bolls and squares from insect damage, than other promoters.

The promoters used in the DNA constructs (i.e. chimeric plant genes) of the present invention may be modified, if desired, to affect their control characteristics. For example, the CaMV35S promoter may be  
30 ligated to the portion of the ssRUBISCO gene that represses the expression of ssRUBISCO in the absence of light, to create a promoter which is active in leaves but not in roots. The resulting chimeric promoter may be used as described herein. For purposes of this description, the phrase "CaMV35S" promoter thus includes variations of CaMV35S promoter,  
35 e.g., promoters derived by means of ligation with operator regions, random

or controlled mutagenesis, etc. Furthermore, the promoters may be altered to contain multiple "enhancer sequences" to assist in elevating gene expression. Examples of such enhancer sequences have been reported by Kay et al. (1987).

5       The RNA produced by a DNA construct of the present invention also contains a 5' non-translated leader sequence. This sequence can be derived from the promoter selected to express the gene, and can be specifically modified so as to increase translation of the mRNA. The 5' non-translated regions can also be obtained from viral RNA's, from  
10 suitable eukaryotic genes, or from a synthetic gene sequence. The present invention is not limited to constructs wherein the non-translated region is derived from the 5' non-translated sequence that accompanies the promoter sequence. As shown below, a plant gene leader sequence which is useful in the present invention is the petunia heat shock protein  
15 70 (Hsp70) leader. [Winter et al.]

As noted above, the 3' non-translated region of the chimeric plant genes of the present invention contains a polyadenylation signal which functions in plants to cause the addition of adenylate nucleotides to the 3' end of the RNA. Examples of preferred 3' regions are (1) the 3' tran-  
20 scribed, non-translated regions containing the polyadenylate signal of *Agrobacterium* tumor-inducing (Ti) plasmid genes, such as the nopaline synthase (NOS) gene and (2) plant genes like the soybean 7s storage protein genes and the pea ssRUBISCO E9 gene. [Fischhoff et al.]

## 25 Plant Transformation and Expression

A chimeric plant gene containing a structural coding sequence of the present invention can be inserted into the genome of a plant by any suitable method. Suitable plant transformation vectors include those derived from a Ti plasmid of *Agrobacterium tumefaciens*, as well as those  
30 disclosed, e.g., by Herrera-Estrella (1983), Bevan (1983), Klee (1985) and EPO publication 0 120 516 (Schilperoort et al.). In addition to plant transformation vectors derived from the Ti or root-inducing (Ri) plasmids of *Agrobacterium*, alternative methods can be used to insert the DNA constructs of this invention into plant cells. Such methods may involve,  
35 for example, the use of liposomes, electroporation, chemicals that increase

free DNA uptake, free DNA delivery via microprojectile bombardment, and transformation using viruses or pollen.

A particularly useful Ti plasmid cassette vector for transformation of dicotyledonous plants is pMON11782. The expression cassette  
5 pMON11782 consists of the FMV35S promoter, the petunia Hsp70 5' untranslated leader, and the 3' end including polyadenylation signals from the pea ssRUBISCO E9 gene. Between the leader and the 3' polyadenylation signals is a multilinker containing multiple restriction sites, including a BamHI site for the insertion of genes. pMON11782 also  
10 contains a HindIII site before the promoter sequence.

The remainder of pMON11782 contains a segment of pBR322 (New England Biolabs, Beverly, MA) which provides an origin of replication in *E. coli*; the oriV region from the broad host range plasmid RK1 which allows replication in *Agrobacterium* strain ABI; the streptomycin-  
15 spectinomycin resistance gene from Tn7; and a chimeric NPTII gene, containing the CaMV35S promoter and the nopaline synthase (NOS) 3' end, which provides kanamycin resistance in transformed plant cells.

Another particularly useful Ti plasmid cassette vector is pMON17227. This vector is described by Barry et al. in WO 92/04449  
20 and contains a gene encoding an enzyme conferring glyphosate resistance which is an excellent selection marker gene for many plants.

#### Transient Expression of 3-Hydroxysteroid Oxidase in Tobacco Plants

Both 3-hydroxysteroid oxidase gene cassettes, that is the gene  
25 encoding intact protein with the signal sequence and that encoding only the mature protein, each modified at the N-terminus as described above, were mobilized as NcoI-BamHI fragments and inserted into a transient expression vector that had been cut with NcoI and BamHI. A transient expression vector is a simple plasmid containing a plant promoter with a  
30 5' nontranslated leader, a 3' nontranslated polyadenylation sequence, and between them a multilinker having multiple restriction sites for insertion of a protein coding sequence. The constructed vectors placed the 3-hydroxysteroid oxidase gene under the control of the FMV35S promoter with the petunia HSP70 leader sequence discussed above. At the 3' end terminator  
35 region is the non-translated polyadenylation signal terminator region of

the nopaline synthase gene. A plasmid containing the intact protein coding sequence (SEQ ID NO:13) was identified and named pMON20910. A plasmid containing the modified mature protein coding sequence (SEQ ID NO:14) was identified and named pMON20908.

5 pMON20910 and pMON20908 are vectors for expression of 3-hydroxysteroid oxidase genes in plant cells, but these vectors lack appropriate sequences for use in *Agrobacterium*-mediated plant transformation. However, these vectors can be used for either transient expression of 3-hydroxysteroid oxidase in plant cells, or they can be used to  
10 generate stably transformed cotton plants via free DNA delivery such as biolistic bombardment of cotton meristems.

For transient expression analysis, plasmid DNA samples from pMON20908 and pMON20910 vectors were purified and introduced into tobacco via electroporation. Freeze-thaw extraction followed by a nine-  
15 fold concentration of soluble fractions on Centricon-10 filter concentrators allowed unambiguous detection of 3-hydroxysteroid oxidase activity in all cell lysates, immunologically by Western blot assay and enzymatically. The activity of the lysate from cells containing pMON20908, that is the coding sequence for the modified mature protein, was approximately ten-  
20 fold lower than that recovered from cells containing pMON20910. Western blot analysis indicated that the signal sequence is cleaved in protoplasts, although not necessarily with the fidelity necessary to generate a processed protein identical in form and activity to that naturally secreted by *Streptomyces* A19249.

25

#### Stable Transformation of Dicots with a 3-Hydroxysteroid Oxidase Gene

pMON20908 and pMON20910 were used to construct vectors for stable transformation of dicots with *Agrobacterium*. Each was partially digested with the restriction enzyme NotI to generate DNA fragments  
30 containing the FMV35S promoter, the petunia HSP70 leader, either the intact [full length] or mature 3-hydroxysteroid oxidase coding sequence, and nos 3' polyadenylation site. Partial digestion was necessary because in addition to NotI sites flanking the expression cassette, there are two NotI sites within the cholesterol oxidase gene - one within the coding region  
35 for the signal peptide and one within the coding region for the mature pro-

tein. The NotI partial digest fragments containing the entire expression cassettes are isolated using agarose gel electrophoresis and purified by extraction from the agarose. These NotI fragments are inserted into NotI-digested plasmid pMON17227, which is described by Barry et al. in  
5 WO 92/04449. pMON20913 was identified as containing the intact coding sequence. pMON20923 was identified as containing the mature coding sequence.

These vectors were introduced into disarmed *Agrobacterium* host ABI and used to transform tobacco in tissue culture. Selection for glypho-  
10 sate resistance led to several transformed lines for each vector. For pMON20913, four out of 18 lines were confirmed as 3-hydroxysteroid oxidase expressors by Western blot assay. For pMON20923, four out of 29 lines were confirmed as 3-hydroxysteroid oxidase expressors by Western blot assay. Analysis of the lines transformed with pMON20913  
15 confirm the presence of the 3-hydroxysteroid oxidase activity. The highest expressing plant had an expression level of 0.2% of total protein, which is equivalent to 54 ng of enzyme per mg of wet tissue.

Vectors containing the intact or mature 3-hydroxysteroid oxidase cassette express the active enzyme in the cytoplasm of the plant cell.  
20 There has been no evidence of secretion outside the transformed cells. Some bacterial secretory signal sequences have been shown to function in plant cells. It may be desirable to direct most or all of the 3-hydroxysteroid oxidase protein into the plant secretory pathway. To achieve this, it may be advantageous to use a signal sequence derived from a plant  
25 gene rather than a bacterial signal. An example of such a signal is that from the tobacco PR1b gene, described by Cornelissen et al. pMON 10824, disclosed in EP Publ. 0 385 962, is a plant transformation vector designed for the expression of the lepidopteran active *B.t. kurstaki* protein. In pMON10824, the *B.t.k.* coding sequence is fused to the PR1b signal  
30 sequence plus 10 amino acids of the mature PR1b coding sequence. This *B.t.k.* fusion gene is driven by the CaMV35S promoter containing a duplicated enhancer. Other vectors carrying the PR1B signal and CaMV35S promoter may also serve as the source of these elements. To create a vector in which the PR1b signal is fused to the 3-hydroxysteroid  
35 oxidase gene, a DNA fragment containing the CaMV35S promoter

sequence and the PR1B signal sequence is used to replace the fragment containing the FMV promoter and HSP70 leader sequences in pMON 20913 and pMON20923. This results in plasmids in which the 3-hydroxysteroid oxidase coding sequence (either mature protein or intact protein cassette) is fused to the amino terminal secretory signal from the PR1b gene and is driven by the CaMV35S promoter. NcoI restriction enzyme sites at the 3' end of the PR1B signal-containing fragment and the 5' end of the modified 3-hydroxysteroid oxidase protein coding sequence allows in-frame translational fusions between the two elements. pMON20930 and pMON20932 are plant transformation vectors that carry such fusions between PR1B signal sequence and either the intact cholesterol oxidase and the mature cholesterol oxidase, respectively. A similar plasmid may be constructed wherein the 3-hydroxysteroid oxidase gene is under the control of the FMV35S promoter. Such plasmids are mobilized into a disarmed *Agrobacterium* host and used to transform dicots. Thus, these plants produce a 3-hydroxysteroid oxidase that is secreted into the extracellular space.

In some cases, proteins that enter the plant secretory pathway are targeted to different cellular compartments such as the vacuole. It may be desirable to direct the 3-hydroxysteroid oxidase to the vacuole of plant cells. In this case, the vectors described above in which the PR1b signal is used are further modified to include vacuolar targeting sequences derived from known plant vacuolar enzyme genes.

It may also be desirable to retain the 3-hydroxysteroid oxidase in the lumen of the endoplasmic reticulum (ER). In this case, vectors in which the PR1b signal is used to target the protein to the secretory pathway are further modified to include sequences known to encode ER retention signals. These sequences are added such that the four-amino acid-long ER retention peptides (such as HDEL (SEQ ID NO:15) and KDEL (SEQ ID NO:16) are fused to the carboxy terminus of the 3-hydroxysteroid oxidase. Site-directed oligonucleotide mutagenesis is used to introduce these additions to the carboxy-terminus coding sequence of 3-hydroxysteroid oxidase. pMON20937 and pMON20938 are vectors in which the peptides HDEL (SEQ ID NO:15) and RGSEKDEL (SEQ ID NO:17) are introduced, respectively, into pMON20932.

It may also be advantageous to direct the localization of the 3-hydroxysteroid oxidase protein to another cellular compartment, the chloroplast. Proteins can be directed to the chloroplast by including at their N-termini a chloroplast transit peptide (CTP). One CTP that has worked to localize heterologous proteins to the chloroplast is that derived from the RUBISCO small subunit gene of *Arabidopsis*, denoted *ats1A*. A variant of this transit peptide that encodes the transit peptide, 24 amino acids of mature RUBISCO sequence, plus a reiteration of the transit peptide cleavage site has been constructed for the successful chloroplast localization of the *B.t.k.* protein (Wong, 1992). Vectors containing the *Arabidopsis* *ats1A* transit peptide fused to the *B.t.k.* gene may be used as the base for constructing vectors for the chloroplast localization of the 3-hydroxysteroid oxidase protein. For example, pMON10817, constructed of the *Arabidopsis* RUBISCO small subunit promoter from the *ats1A* gene, the native *ats1A* 5' untranslated leader, the *ats1A* chloroplast transit peptide variant described above and a truncated *B.t.k.* enzyme gene is digested with restriction enzymes NotI and NcoI. A NotI - NcoI DNA fragment containing the promoter, leader and transit peptide is used to replace a NotI - NcoI fragment carrying the FMV promoter and HSP70 leader sequences in pMON20913 and pMON20923. These reactions construct plasmids pMON20929 and pMON20931 in which the 3-hydroxysteroid oxidase coding sequence (either intact protein or mature protein cassette, respectively) is fused at its amino terminus to the chloroplast transit peptide and is transcribed from the *ats1A* promoter. Alternatively, a similar plasmid may be constructed to replace the *ats1A* promoter with the CaMV35S or FMV35S promoters. Such plasmids are mobilized into disarmed *Agrobacterium* hosts and used to transform dicots. Thus, these plants produce a 3-hydroxysteroid oxidase that is localized to the chloroplast.

It may also be advantageous to direct the 3-hydroxysteroid oxidase to another subcellular compartment, the mitochondria. Examples of proteins that are normally localized to the mitochondria and for which the targeting sequences are known include cytochrome C1 which is localized to the mitochondrial intermembrane space (Braun, 1992) and the  $\beta$  subunit of ATP synthase which is localized to the mitochondrial matrix



(Boutry, 1987). Mitochondrial targeting DNA sequences encoding the amino terminal mitochondrial targeting peptides are isolated using the polymerase chain reaction (PCR). Oligonucleotide primers corresponding to the amino and carboxy terminal portions of these targeting sequences  
5 are used to PCR-amplify first strand cDNA that has been generated from mRNA using reverse transcriptase. The oligonucleotide primers have NcoI restriction enzyme sites attached so that the amplified product, when cleaved with NcoI, can be inserted into the NcoI sites at the amino terminus of the cloned 3-hydroxysteroid oxidase gene cassettes. Thus,  
10 various forms of 3-hydroxysteroid oxidase fused to mitochondrial targeting peptides can be expressed in plants using promoters such as CaMV35S or FMV35S, and can be localized to the mitochondria.

#### Stable Transformation of Monocots

15 A 3-hydroxysteroid oxidase gene may be stably incorporated into the genome of monocots using the vectors and methods described in WO 93/19189 (Brown et al.). The gene can be inserted in an appropriate vector, for example pMON19653 and pMON19643, described by Brown et al. The resulting construct contains a cassette of the CaMV E35S  
20 promoter, the Hsp70 intron, the CP4 glyphosate selection marker, and the NOS terminator; a cassette of the CaMV E35S promoter, the Hsp70 intron, the GOX glyphosate selection marker, and the NOS terminator; and a single NotI site for insertion of a gene expression cassette containing a 3-hydroxysteroid oxidase gene, such as SEQ ID NO:13 or  
25 SEQ ID NO:14.

This vector is inserted by bombardment of embryogenic tissue culture cells using a biolistic particle gun as described by Brown et al. Transformed cells are selected for glyphosate resistance and whole plants are regenerated. Insect-resistant plants may be confirmed to be  
30 expressing the gene by Western blot analysis, esterase activity assay, or insect resistance assay.

Targeting of the protein to certain cellular compartments is also possible in monocots using the signal sequences described above.

From the foregoing, it will be seen that this invention is one well  
35 adapted to attain all the ends and objects hereinabove set forth together

with advantages which are obvious and which are inherent to the invention. It will be understood that certain features and subcombinations are of utility and may be employed without reference to other features and subcombinations. This is contemplated by and is within the scope of the

5 claims. Since many possible embodiments may be made of the invention without departing from the scope thereof, it is to be understood that all matter herein set forth or shown in the accompanying drawings is to be interpreted as illustrative and not in a limiting sense.

REFERENCES

- Bagdasarian, M., et al. Gene, 16: 237-47, 1981.
- Bevan, M. et al., Nature, 304:184, 1983.
- Boutry, et al., Nature 328: 340-342, 1987.
- 5 Braun, et al., Mol. Gen. Genet. 231: 217-225, 1992.
- Burnette, W.N. Anal. Biochem., 112: 195-203, 1981.
- Cornelissen, B.J.C., et al. EMBO Journal, 5: 37-40, 1986.
- Fischhoff, D.A. and Perlak, F.J. "Synthetic plant genes and method for preparation." European Patent Application, Publication Number 0  
10 385 962, 1990.
- Gallo, L.L. Methods Enzymol., 71: 665-7, 1981.
- Herrera-Estrella, L. et al., Nature, 303:209, 1983.
- Ishizaki, T., et al. Journal of Bacteriology, 171: 596-601, 1989.
- Kay, R. et al., Science, 236: 1299-1302, 1987.
- 15 Klee, H. J. et al., Bio/Technology, 3: 637-642, 1985.
- Knauf, V.C. and Nester, E. Plasmid, 8: 43-54, 1982.
- Laemmli, U.K. Nature, 227: 680-5, 1970.
- Long, Susan, and Ostroff, Gary R. PCT Int. Appl. WO 90 05,788.
- Marrone, P.G., et al. J. Econ. Entom., 78: 290-3, 1985.
- 20 Matsudaira, P. J. Biol. Chem., 261: 10035-38, 1987.
- Moore, S. and Stein, W.H. Methods in Enzymology, 6: 819-31, 1963.
- Purcell, J.P., Greenplate, J.T., and Sammons, R.P. Insect Biochem. Molec. Biol., 22:41-47, 1992.
- Purcell, J.P., et al. Biochem. Biophys. Res. Comm. 196:1406-1413, 1993.
- 25 Schuler, M. A. et al., Nucleic Acids Research, 10: 8225-8244, 1982.
- Smith, A. G., and Brooks, C.J.W. J. Steroid Biochem., 7: 705-713, 1976.
- Smith, P.K., et al. Analytical Biochemistry, 150: 76-85, 1985.
- Winter et al. Mol. Gen. Genet., 221(2): 315-19, 1988.
- Wong, et al., Plant Mol. Biol. 20:81-93, 1992.
- 30 Yanisch-Perron, C., et al. Gene, 33: 103-19, 1985.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- 5 (i) APPLICANT:  
(A) NAME: Monsanto Company  
(B) STREET: 800 North Lindbergh Boulevard  
(C) CITY: St. Louis  
(D) STATE: Missouri  
10 (E) COUNTRY: United States of America  
(F) POSTAL CODE (ZIP): 63167  
(G) TELEPHONE: (314)694-3131  
(H) TELEFAX: (314)694-5435
- 15 (ii) TITLE OF INVENTION: Method of Controlling Insects  
(iii) NUMBER OF SEQUENCES: 14
- 20 (iv) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
- 25 (vi) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 07/762682  
(B) FILING DATE: 23-SEP-1991
- 30 (vi) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 07/937195  
(B) FILING DATE: 09-SEP-1992
- 35 (vi) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 08/083948  
(B) FILING DATE: 28-JUN-1993

## (2) INFORMATION FOR SEQ ID NO:1:

- 40 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear
- 45 (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Ser | Thr | Leu | Met | Leu | Glu | Met | Gly | Gln | Leu | Trp | Asn | Gln | Pro |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     | 15  |     |

## (2) INFORMATION FOR SEQ ID NO:2:

- 55 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear
- 60 (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ala Phe Ala Asp Asp Phe Cys Tyr His Pro Leu Gly Gly Cys Val Leu  
1 5 10 15

5

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Asn Leu Tyr Val Thr Asp Gly Ser Leu Ile Pro Gly  
1 5 10

20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTGTCCACCC TGATGCTGGA GATGGGCCAG CTGTGGAACC AGCCC

35

45

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 48 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCCTTCGCCG ACGACTTCTG CTACCACCCG CTCGGCGGCT GCGTCCTG

50

48

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

55

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AACCTCTACG TGACCGACGG TTCGCTGATC CCGGGT

36

5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1865 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTACTCCATG GCGTGCTGAA GGTGCGTGCC TGGCCTCCCG AGGTCGTCGA GGACTTCGTG 60  
20 AAGTGAGCGG GCACCCCGCC CGTCCCGCC CCGCAACGGC CCGTTCGCA CACCGGGTGA 120  
CCCGACCCCG TCGGCCCCCG ACGTCCGCG ACCTCTCAGT CCCCTCTCGA AGCTCAGSAG 180  
25 CAACAGCGTG AACGCACACC AGCCTCTGTC GCGCCGCGGC ATGCTCGGCC TGGCCGCCTT 240  
GGGCGCCGCC GCACTCACC GGCAGACCAC GATCACC GCGCCCGCG CGGCCGCGGC 300  
CACCGCCCC GCGGCTCCG GCGGCACGTT CGTGCCCGCC GTCGTGATCG GCACCGGCTA 360  
30 CGGCGCGGCC GTCTCCGCC TCGGCTCGG CGAGGCCGGG GTCTCCACCC TGATGCTGGA 420  
GATGGGCCAG CTGTGGAACC AGCCCGGCC GGACGGCAAC GTCTCTGCG GGATGCTCAA 480  
GCCCCACAAG CGTCCAGCT GGTCAAGAC CCGCACCGAG GCCCCGCTCG GCTCCTTCCT 540  
35 CTGGCTCGAC CTCGCCAACC GGGACATCGA CCCCTACGCG GCGCTCCTGG ACCGGGTCAA 600  
CTTCGACCAG ATGTCCGTGT ACGTGGGCG CGGGGTGCG GCGGCTCGC TCGTCAACGG 660  
40 CGGTATGGCC GTCACGCCCC GCGCTCCTA CTTCCAGGAG GTGCTGCCCC AGGTCGACGC 720  
CGACGAGATG TACGGCACCT ACTTCCCGCG CGCGAATCC GGCCTGCGGG TCAACAACAT 780  
CGACAAGGAC TGGTTCGAGC AGACCGAGTG GTACACGTT CCGCGCGTTG CCCGTCTGCA 840  
45 GGGCGAGAAC GCGGCTCGA AGACCACCTT CGTGCCCAAC GTCTACGACT GGGACTACAT 900  
GCGCGGTGAG GCGGACGGCA CCAACCCCAA GTCCGCGCTC GCGCCGAGG TCATCTACGG 960  
50 CAACAACCAC GGCAAGGTCT CCCTCGACAA GAGCTACCTG GCGGCCGCCC TGGGCACCGG 1020  
CAAGGTCACC GTCGAGACCC TGCACCAGGT CAAGACGATC CGTCAGCAGA ACGACGGCAC 1080  
CTACCTGCTG ACGGTCGAGC AGAAGGACCC CGACGGCAAG CTGCTCGGGA CCAAGGAGAT 1140  
55 CTCTGCGCG CACTCTTCC TCGGCGCCGG CAGCCTCGGC TCCATTGAAC TGCTGCTGCG 1200  
CGCCCGGGAG ACCGGCACCC TGCCCGGCT CAGCTCCGAG ATCGGCGGCG GCTGGGGCCC 1260  
60 CAACGGCAAC ATCATGACCG CCCGCGCCAA CCATGTGTGG AACCCACGG GCAGCAAGCA 1320  
GTCGTCGATC CCCGCCCTCG GCATCGACGA CTGGGACAAC CCCGACAACC CCGTCTTCGC 1380

CGAGATAGCC CCCATGCCGG CGGGCCTCGA GACCTGGGTC AGCCTCTACC TGGCCATCAC 1440  
 CAAGAACCCG GAGCGCGGCA CCTTCGTCTA CGACGCCGCC AAGGACCGGG CGGACCTGCG 1500  
 5 CTGGACCCGG GACCAGAACG CGCCCGCGGT CGCCGCCGCC AAGTCGCTGT TCGACCGCGT 1560  
 CAACAAGGCC AACACGACCA TCTACCGGTA CGACCTCTTC GGCAAGCAGA TCAAGGCGTT 1620  
 10 CGCCGACGAC TTCTGCTACC ACCCGCTCGG CGGCTGCGTC CTCGGCAAGG CCACCGACAA 1680  
 CTACGGCCGC GTCTCCGGGT ACAAGAACCT CTACGTCACC GACGGCTCGC TCATCCCCGG 1740  
 CAGCATCGGC GTCAACCCGT TCGTGACCAT CACGGCGCTG GCGGAGCGGA ACGTCGAGCG 1800  
 15 CGTCATCAAG GAGGACATCG CGGGTTCTTG ACGAGCGACG GGCGGGGCGC GGCATGCAAG 1860  
 CTTGG 1865

20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 547 amino acids  
 25 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Val Asn Ala His Gln Pro Leu Ser Arg Arg Arg Met Leu Gly Leu Ala  
 1 5 10 15  
 35 Ala Leu Gly Ala Ala Ala Leu Thr Gly Gln Thr Thr Ile Thr Ala Ala  
 20 25 30  
 Pro Arg Ala Ala Ala Ala Thr Ala Pro Gly Gly Ser Gly Gly Thr Phe  
 35 40 45  
 40 Val Pro Ala Val Val Ile Gly Thr Gly Tyr Gly Ala Ala Val Ser Ala  
 50 55 60  
 45 Leu Arg Leu Gly Glu Ala Gly Val Ser Thr Leu Met Leu Glu Met Gly  
 65 70 75 80  
 Gln Leu Trp Asn Gln Pro Gly Pro Asp Gly Asn Val Phe Cys Gly Met  
 85 90 95  
 50 Leu Lys Pro Asp Lys Arg Ser Ser Trp Phe Lys Thr Arg Thr Glu Ala  
 100 105 110  
 Pro Leu Gly Ser Phe Leu Trp Leu Asp Leu Ala Asn Arg Asp Ile Asp  
 115 120 125  
 55 Pro Tyr Ala Gly Val Leu Asp Arg Val Asn Phe Asp Gln Met Ser Val  
 130 135 140  
 60 Tyr Val Gly Arg Gly Val Gly Gly Gly Ser Leu Val Asn Gly Gly Met  
 145 150 155 160  
 Ala Val Thr Pro Arg Arg Ser Tyr Phe Gln Glu Val Leu Pro Gln Val

	Asp	Ala	Asp	Glu	Met	Tyr	Gly	Thr	Tyr	Phe	Pro	Arg	Ala	Asn	Ser	Gly	
					180					185						190	
5																	
	Leu	Arg	Val	Asn	Asn	Ile	Asp	Lys	Asp	Trp	Phe	Glu	Gln	Thr	Glu	Trp	
			195					200					205				
10	Tyr	Thr	Phe	Ala	Arg	Val	Ala	Arg	Leu	Gln	Ala	Glu	Asn	Ala	Gly	Leu	
		210					215					220					
	Lys	Thr	Thr	Phe	Val	Pro	Asn	Val	Tyr	Asp	Trp	Asp	Tyr	Met	Arg	Gly	
	225					230					235					240	
15	Glu	Ala	Asp	Gly	Thr	Asn	Pro	Lys	Ser	Ala	Leu	Ala	Ala	Glu	Val	Ile	
					245					250					255		
	Tyr	Gly	Asn	Asn	His	Gly	Lys	Val	Ser	Leu	Asp	Lys	Ser	Tyr	Leu	Ala	
			260						265					270			
20																	
	Ala	Ala	Leu	Gly	Thr	Gly	Lys	Val	Thr	Val	Glu	Thr	Leu	His	Gln	Val	
			275					280					285				
	Lys	Thr	Ile	Arg	Gln	Gln	Asn	Asp	Gly	Thr	Tyr	Leu	Leu	Thr	Val	Glu	
25		290					295					300					
	Gln	Lys	Asp	Pro	Asp	Gly	Lys	Leu	Leu	Gly	Thr	Lys	Glu	Ile	Ser	Cys	
	305					310					315					320	
30	Arg	His	Leu	Phe	Leu	Gly	Ala	Gly	Ser	Leu	Gly	Ser	Ile	Glu	Leu	Leu	
					325					330					335		
	Leu	Arg	Ala	Arg	Glu	Thr	Gly	Thr	Leu	Pro	Gly	Leu	Ser	Ser	Glu	Ile	
				340					345					350			
35																	
	Gly	Gly	Gly	Trp	Gly	Pro	Asn	Gly	Asn	Ile	Met	Thr	Ala	Arg	Ala	Asn	
		355						360					365				
	His	Val	Trp	Asn	Pro	Thr	Gly	Ser	Lys	Gln	Ser	Ser	Ile	Pro	Ala	Leu	
40		370					375					380					
	Gly	Ile	Asp	Asp	Trp	Asp	Asn	Pro	Asp	Asn	Pro	Val	Phe	Ala	Glu	Ile	
	385					390					395					400	
45	Ala	Pro	Met	Pro	Ala	Gly	Leu	Glu	Thr	Trp	Val	Ser</					



32

Asp Asn Tyr Gly Arg Val Ser Gly Tyr Lys Asn Leu Tyr Val Thr Asp  
500 505 510

5 Gly Ser Leu Ile Pro Gly Ser Ile Gly Val Asn Pro Phe Val Thr Ile  
515 520 525

Thr Ala Leu Ala Glu Arg Asn Val Glu Arg Val Ile Lys Glu Asp Ile  
530 535 540

10 Ala Gly Ser  
545

(2) INFORMATION FOR SEQ ID NO:9:

15

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

25

CTCAGGAGCA CCATGGCGAC CGCACAC

27

(2) INFORMATION FOR SEQ ID NO:10:

30

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

40

GTGCCGCCCG AGGCCATGGG GCGGTGGC

29

(2) INFORMATION FOR SEQ ID NO:11:

45

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 30 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

55

GCCCCGCCCG TCGGATCCGT CAGGAACCCG

30

(2) INFORMATION FOR SEQ ID NO:12:

60

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ser Gly Gly Thr Phe

1

5

10

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 1647 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

25 ATGGCGACCG CACACCAGCC TCTGTCGCGC CGCCGCATGC TCGGCCTGGC CGCCTTGGGC 60  
GCCGCCGCAC TCACCGGGCA GACCACGATC ACCCGGGCCC CCCGCGCGGC CGCCGCCACC 120  
GCCCCCGGCG GCTCCGGCGG CACGTTCGTG CCCGCCGTG TGATCGGCAC CGGCTACGGC 180  
30 CGCGCCGTCT CGCCCTGCG GCTCGGCGAG GCCGGGTCT CCACCCTGAT GCTGGAGATG 240  
GGCCAGCTGT GGAACCAGCC CGGCCCGGAC GGCAACGTCT TCTGCGGGAT GCTCAAGCCC 300  
GACAAGCGCT CCAGCTGGTT CAAGACCCGC ACCGAGGCC CGCTCGGCTC CTTCTCTGG 360  
35 CTCGACCTCG CCAACCGGA CATCGACCCC TACGCGGGCG TCCTGGACCG GGTCAACTTC 420  
GACCAGATGT CCGTGTACGT GGGCCGCGG GTCGGCGGCG GCTCGCTCGT CAACGGCGGT 480  
40 ATGGCCGTCA CGCCCCGGCG CTCCTACTTC CAGGAGGTGC TGCCCCAGGT CGACGCCGAC 540  
GAGATGTACG GCACCTACTT CCCGCGCGCG AACTCCGGCC TCGGGGTCAA CAACATCGAC 600  
AAGGACTGGT TCGAGCAGAC CGAGTGGTAC ACGTTCGCGC GCGTTGCCCG TCTGCAGGCC 660  
45 GAGAACGCCG GCCTGAAGAC CACCTTCGTG CCCAACGTCT ACGACTGGGA CTACATGCGC 720  
GGTGAGGCGG ACGGCACCAA CCCCAAGTCC GCGCTCGCG CCGAGGTCAT CTACGGCAAC 780  
50 AACCACGGCA AGGTCTCCCT CGACAAGAGC TACCTGGCGG CGCCCTGGG CACCGGCAAG 840  
GTCACCGTCG AGACCTGCA CCAGGTCAAG ACGATCCGTC AGCAGAACGA CGGCACCTAC 900  
CTGCTGACGG TCGAGCAGAA GGACCCCGAC GGCAAGCTGC TCGGGACCAA GGAGATCTCC 960  
55 TGCCGCCACC TCTTCTCGG CGCCGGCAGC CTCGGCTCCA TTGAAGTCT GCTGCGCGCC 1020  
CGGGAGACCG GCACCTGCG CGGCCTCAGC TCCGAGATCG GCGGCGGCTG GGGCCCCAAC 1080  
60 GGCAACATCA TGACCGCCCG CGCCAACCAT GTGTGGAACC CCACGGGCAG CAAGCAGTCG 1140  
TCGATCCCCG CCCTCGGCAT CGACGACTGG GACAACCCCG ACAACCCCGT CTTGCGCGAG 1200

ATAGCCCCCA TGCCGGCGGG CCTCGAGACC TGGGTCAGCC TCTACCTGGC CATCACCAAG 1260  
AACCCGGAGC GCGGCACCTT CGTCTACGAC GCCGCCAAGG ACCGGGCGGA CCTGCGCTGG 1320  
5 ACCCGGGACC AGAACGCGCC CGCGGTGCGC GCCGCCAAGT CGCTGTTCGA CCGCGTCAAC 1380  
AAGGCCAACA CGACCATCTA CCGGTACGAC CTCTTCGGCA AGCAGATCAA GGCCTTCGCC 1440  
10 GACGACTTCT GCTACCACCC GCTCGGGCGG TGCCTCCTCG GCAAGGCCAC CGACAACCTAC 1500  
GGCCGCGTCT CCGGTACAA GAACCTCTAC GTCACCGACG GCTCGCTCAT CCCCGGCAGC 1560  
ATCGGCGTCA ACCCGTTCGT GACCATCAGC GCGCTGGCGG AGCGGAACGT CGAGCGCGTC 1620  
15 ATCAAGGAGG ACATCGCGGG TTCCTGA 1647

## (2) INFORMATION FOR SEQ ID NO:14:

20

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1521 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

25

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

30

ATGGCCTCCG GCGGCACGTT CGTCCCCGCC GTCGTGATCG GCACCGGCTA CGGCGCGGCC 60  
GTCTCCGCCC TGCGGCTCGG CGAGGCGGG GTCTCCACCC TGATGCTGGA GATGGGCCAG 120  
35 CTGTGGAACC AGCCCCGCCC GGACGGCAAC GTCTTCTGCG GGATGCTCAA GCCGACAAG 180  
CGCTCCAGCT GGTTCAGAC CCGCACCGAG GCCCGCTCG GCTCCTTCCT CTGGCTCGAC 240  
CTCGCCAACC GGGACATCGA CCCCTACGCG GCGCTCCTGG ACCGGGTCAA CTTGACCAG 300  
40 ATGTCCGTGT ACGTGGGCCG CGGGGTGGC GCGGCTCGC TCGTCAACGG CGGTATGGCC 360  
GTCACGCCCC GCGCTCCTA CTTCAGGAG GTGTGCCCC AGGTCGACGC CGACGAGATG 420  
45 TACGGCACCT ACTTCCCGC CGCGAACTCC GGCCTGCGGG TCAACAACAT CGACAAGGAC 480  
TGGTTCGAGC AGACCGAGTG GTACAGTTC GCGCGCGTTG CCCGTCTGCA GGCCGAGAAC 540  
GCCGGCCTGA AGACCACCTT CGTGCCCAAC GTCTACGACT GGGACTACAT GCGCGGTGAG 600  
50 GCGGACGGCA CCAACCCCAA GTCCGCGCTC GCCGCCGAGG TCATCTACGG CAACAACCAC 660  
GGCAAGGTCT CCCTCGACAA GAGCTACCTG GCGGCCGCCC TGGGCACCGG CAAGGTCACC 720  
55 GTCGAGACCC TGCACCAGGT CAAGACGATC CGTCAGCAGA ACGACGGCAC CTACCTGCTG 780  
ACGGTCGAGC AGAAGGACCC CGACGGCAAG CTGCTCGGGA CCAAGGAGAT CTCCTGCCGC 840  
CACCTCTTCC TCGGCGCCGG CAGCCTCGGC TCCATTGAAC TGCTGCTGCG CGCCCGGGAG 900  
60 ACCGGCACCC TGCCCGGCT CAGCTCCGAG ATCGGCGGCG GCTGGGGCCC CAACGGCAAC 960

ATCATGACCG CCCGCGCCAA CCATGTGTGG AACCCACGG GCAGCAAGCA GTCGTCGATC 1020  
CCCGCCCTCG GCATCGACGA CTGGGACAAC CCCGACAACC CCGTCTTCGC CGAGATAGCC 1080  
5 CCCATGCCGG CGGGCCTCGA GACCTGGGTC AGCCTCTACC TGGCCATCAC CAAGAACCCG 1140  
GAGCGCGGCA CCTTCGTCTA CGACGCCGCC AAGGACCGGG CGGACCTGCG CTGGACCCGG 1200  
GACCAGAACG CGCCCGCGGT CGCCGCCGCC AAGTCGCTGT TCGACCGCGT CAACAAGGCC 1260  
10 AACACGACCA TCTACCGGTA CGACCTCTTC GGCAAGCAGA TCAAGGCGTT CGCCGACGAC 1320  
TTCGTCTACC ACCCGCTCGG CGGCTGCGTC CTCGGCAAGG CCACCGACAA CTACGGCCGC 1380  
15 GTCTCCGGGT ACAAGAACCT CTACGTCACC GACGGCTCGC TCATCCCCGG CAGCATCGGC 1440  
GTCAACCCGT TCGTGACCAT CACGGCGCTG GCGGAGCGGA ACGTCGAGCG CGTCATCAAG 1500  
GAGGACATCG CGGGTTCCTG A 1521  
20

(2) INFORMATION FOR SEQ ID NO:15:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 4 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear
- 30 (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:
- His Asp Glu Leu  
1

35

(2) INFORMATION FOR SEQ ID NO:16:

- 40 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 4 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
- Lys Asp Glu Leu  
1

50 (2) INFORMATION FOR SEQ ID NO:17:

- 55 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 8 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
- Arg Gly Ser Glu Lys Asp Glu Leu  
1 5

## WHAT IS CLAIMED IS:

1. A method of controlling lepidopteran insect infestation of plants comprising providing a 3-hydroxysteroid oxidase for ingestion by the insect.
- 5 2. The method of Claim 1 wherein the insect is in a larval stage.
3. The method of Claim 1 wherein said 3-hydroxysteroid oxidase is provided by plant-colonizing microorganisms which produce 3-hydroxysteroid oxidase after application to the plant.
4. The method of Claim 1 wherein said 3-hydroxysteroid oxidase is  
10 provided by expression of a gene for 3-hydroxysteroid oxidase incorporated in the plant by previous genetic transformation of a parent cell of the plant.
5. The method of Claim 4 wherein said plant is cotton or corn.
6. A method of producing a genetically transformed plant which  
15 expresses an amount of a 3-hydroxysteroid oxidase effective to control lepidopteran insect infestation, comprising the steps of:
  - a) inserting into the genome of a plant cell a recombinant, double-stranded DNA molecule comprising
    - (i) a promoter which functions in plant cells to cause the  
20 production of an RNA sequence;
    - (ii) a structural coding sequence that encodes for 3-hydroxysteroid oxidase;
    - (iii) a 3' non-translated region which functions in said plant cells to cause the addition of polyadenylate nucleotides to the  
25 3' end of the RNA sequence,wherein said promoter is heterologous with respect to said structural coding sequence and wherein said promoter is operatively linked with said structural coding sequence, which is in turn operably linked with said non-translated region;
  - 30 b) obtaining transformed plant cells; and
  - c) regenerating from the transformed plant cells genetically transformed plants with express an amount of 3-hydroxysteroid oxidase effective to control lepidopteran insect infestation.
7. The method of Claim 6 wherein said structural DNA sequence  
35 comprises SEQ ID NO:13 or SEQ ID NO: 14.

8. The method of Claim 6 wherein said plant cell is a cotton or corn plant cell.
9. The method of Claim 6 wherein wherein the genome of said plant cell also contains one or more genes expressing *B.t.* endotoxins.

ID AF263912 standard; DNA; PRO; 123580 BP.  
 XX  
 AC AF263912;  
 XX  
 SV AF263912.1  
 XX  
 DT 25-MAY-2000 (Rel. 63, Created)  
 DT 25-MAY-2000 (Rel. 63, Last updated, Version 1)  
 XX  
 DE Streptomyces noursei ATCC 11455 nystatin biosynthetic gene cluster,  
 DE complete sequence.  
 XX  
 KW .  
 XX  
 OS Streptomyces noursei  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 XX  
 RN [1]  
 RP 1-123580  
 RA Brautaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,  
 RA Valla S., Zotchev S.B.;  
 RT "Biosynthesis of the polyene antifungal antibiotic nystatin in Streptomyces  
 RT noursei ATCC 11455: analysis of the gene cluster and deduction of the  
 RT biosynthetic pathway";  
 RL Chem. Biol. 7(6):395-403(2000).  
 XX  
 RN [2]  
 RP 1-123580  
 RA Brautaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,  
 RA Valla S., Zotchev S.B.;  
 RT ;  
 RL Submitted (04-MAY-2000) to the RL Unigen, NTNU, O. Kyrres gt. 3, Trondheim N-  
 XX  
 DR SPTREMBL; Q9L4V6; Q9L4V6.  
 DR SPTREMBL; Q9L4V7; Q9L4V7.  
 DR SPTREMBL; Q9L4V8; Q9L4V8.  
 DR SPTREMBL; Q9L4V9; Q9L4V9.  
 DR SPTREMBL; Q9L4W0; Q9L4W0.  
 DR SPTREMBL; Q9L4W1; Q9L4W1.  
 DR SPTREMBL; Q9L4W2; Q9L4W2.  
 DR SPTREMBL; Q9L4W3; Q9L4W3.  
 DR SPTREMBL; Q9L4W4; Q9L4W4.  
 DR SPTREMBL; Q9L4W5; Q9L4W5.  
 DR SPTREMBL; Q9L4W6; Q9L4W6.  
 DR SPTREMBL; Q9L4W7; Q9L4W7.  
 DR SPTREMBL; Q9L4W8; Q9L4W8.  
 DR SPTREMBL; Q9L4W9; Q9L4W9.  
 DR SPTREMBL; Q9L4X0; Q9L4X0.  
 DR SPTREMBL; Q9L4X1; Q9L4X1.  
 DR SPTREMBL; Q9L4X2; Q9L4X2.  
 DR SPTREMBL; Q9L4X3; Q9L4X3.  
 DR SPTREMBL; Q9L4X4; Q9L4X4.  
 DR SPTREMBL; Q9L4X5; Q9L4X5.  
 DR SPTREMBL; Q9L4X6; Q9L4X6.  
 DR SPTREMBL; Q9L4X7; Q9L4X7.  
 XX  
 FH Key Location/Qualifiers  
 FH  
 FT source 1..123580

```

FT      /db_xref="taxon:1971"
FT      /organism="Streptomyces noursei"
FT      /strain="ATCC 11455"
FT  CDS  complement(46..783)
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4X7"
FT      /note="putative 4'-phosphopantheteine transferase"
FT      /transl_table=11
FT      /gene="nysF"
FT      /function="presumably post-translationally modifies ACP
FT      domains on PKS"
FT      /product="NysF"
FT      /protein_id="AAF71762.1"
FT      /translation="MIELILPATVATEAAYDDRPRPGDRLLSSEREVIARAVESRQREF
FT      TTVRHLARRALRRLGHPDRAILPNRRGAPQWPPGIVGSMTHCAGYRAAAVSPAELSAAV
FT      SIDAEPNGPLPAGVLNAIALPSERPHLVALAAHRPDVHWRLLFSAKESVFKAWYPLTQ
FT      RELDFSEAEIVIDPTQGAFTARLLVPGPLLGGRRVTVFPGRWHSTPALLTTAVHLPAPT
FT      PRRDREHRTHLTVNSPLPRPTFG"
FT  CDS  complement(867..2684)
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4X6"
FT      /note="putative transporter (ABC family)"
FT      /transl_table=11
FT      /gene="nysG"
FT      /function="presumably involved in efflux of nystatin"
FT      /product="NysG"
FT      /protein_id="AAF71763.1"
FT      /translation="MASPDDLEEERTAPRPARRLVGLLRPHRRSVALAVSMGVGGIVLN
FT      AFGPLLLGRVTDLIADGVLGGVPGPAPGIDFAAIGRLLLVLALYVVASLFMLAQGRLV
FT      ASAVWRTIHELRRDAREKLTRLPLRHFDRQPAGELLSTRTTNDIDNLOQTLLQTLAELIT
FT      SIFSLTMLVLMVLVISPLAVVMLLSVPVSALIAARISKRAQPHYAAQWSANGTLNAHV
FT      EEVCTGHALIKGFDRRAAAEERFDACNDAVYRAAAKAQFASGAMEPVMMFVANLGYVLV
FT      AVIGAWKIVINGTLTLGDVQAFILYARQFSQPIVEIASVAGRLQSGIASAQRVFTLLDAP
FT      EQAPDPLRPGTPARAEGRVEFTDVSFRYSPDTPLIENLSLTVEPGSTVAIVGPTGAGKT
FT      TLGNLLMRFYEPDSGRILLDGTDTATMTRDDLRSRFGVLVLDQTLWFGGTIAENIAYGAP
FT      GACRADIEEAARATCADRFIRTLPGQYDTVLDDESGTVSAGEKQLLTVARAFLARPAVL
FT      VLDEATSSVDTRTEVLIQRAMNSLRAGRTSFVIAHRLSTIRDADLIVMDAGRIVEQGT
FT      HDQLLCAQGLYARLHAARTHTPTAGAAAG"
FT  CDS  complement(2662..4416)
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4X5"
FT      /note="putative transporter (ABC family)"
FT      /transl_table=11
FT      /gene="nysH"
FT      /function="presumably involved in efflux of nystatin"
FT      /product="NysH"
FT      /protein_id="AAF71764.1"
FT      /translation="MLLRLLRAQLRPYAWATAALVALQLVQILGTLTLLPTLGAALIDQG
FT      VVRGDGGRITELGVVMGVVALVQIAAALGAAALAARTATAMGRDLRSALFRRILDFSAR
FT      EIGRFGTPSLLTRSVDVQVQNLAAQTGFGIVVCAPLMCLGSVLLALRQDVPLALLLVA
FT      LVLVAVCFGLLLARMGTLYARMQLTLDRLLGRLLREAITGVRVRSFVRDDHERARFAQ
FT      TND AFLVVSRRVGRLIATMLPVVLLLMNGFTVALLWTGSHRIDAGRMPIGSLSALLSYL
FT      SLILMSVVMLAFVFLSVPRARVCAGRIAEVLDTGSSVAPPAAPQPVGRGPAGRIELCAAG
FT      YRYPGAEEPVL RDVLTVEPGERIAVLGSTGSGKTTLLNLVLRADATEGAVRVGGTDV
FT      RELTAATLAAAVGFVPQRPYLFSGTVASNLRFGRPDATDEELWEALRVAQAADFVARM
FT      DGLDAEITQGGGNVSGGQRQLSLARALLRRPEIYLFDDCF SALDQATDAALRTALVPY
FT      TAGATVITVAQRISAGRDADRIVVLDRGRVVAQGTHTDVLRLTSPTYREIALSQLTEEEA
FT      AHGLAGRP"
FT  CDS  4714..5748
FT      /codon_start=1

```



```

FT      /db_xref="SPTREMBL:Q9L4X4"
FT      /note="putative dGDP-mannose-4,6-dehydratase"
FT      /transl_table=11
FT      /gene="nysDIII"
FT      /function="presumably involved in mycosamine biosynthesis"
FT      /product="NysDIII"
FT      /protein_id="AAF71765.1"
FT      /translation="MSKRALITGITGQDGSYLAEHLLSQGYQVWGLIRGQANPRKSRVS
FT      RLASELDFIDGLMDQGSLSVSAVDTVQPDEVYNLGAISFVPMWSWQQAELVTEVNGMGVL
FT      RMLEAIRMVSGLSSTSRVSPRGQIRFYQASSEMFGKAAETPQRETTLFHPRSPYGAAG
FT      AYGHYITRNYRESFGMYAVSGMLFNHESPRRGQEFVTRKISLAVARIKQGLQDKLALGN
FT      LDAVRDWGYAGDYVRAMHMLQQDAGDDYVIGTGQMHSVRDAVRIAFEHVGLNWDYV
FT      IDPDLVRPAEVEVLCADSAKAQDRLGWKPDVDFPTLMRMMVSDLAQVSRENQYGDVLL
FT      AANW"
FT      CDS      5930..34363
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4X3"
FT      /note="polyketide synthase"
FT      /transl_table=11
FT      /gene="nysI"
FT      /function="responsible for condensation steps 9 to 14 in
FT      the nystatin polyketide backbone synthesis"
FT      /product="NysI"
FT      /protein_id="AAF71766.1"
FT      /translation="MDNEQKLRDYLKLATADLRRTRRRVHKLESAAQEPVAIIGMTCRY
FT      PGGVRSPEDLWRMVEAGEHGVTPFPTDRGWDLEALAAAPTASGGFLHDAPDFDADFFGI
FT      SPREAVAMDPQQRVLESAAWEAFERAGIDPTSVKGSRTGVFIGAMADYRVGPADGAEG
FT      FQLTGNTGSVLSGRISYTFGTGVPVAVTDTACSSSLVAVHLATQALRAGECTLALAGGV
FT      TIMSGPGTFIEMGRQGGLSADGRCRSFGDTADGTGWAEGVGILVLERLSDAVRNGHEIL
FT      AVVRGTAVNQDGASNGLTAPNGPSQQQVIQQALVNARLAAGDIDVVEAHGTGTTLGDPV
FT      EAQALLATYQGNRPADRPLLLGSVKSNSHTQAAAGVAGVIKVMAMRHGTLPRTLHAE
FT      EPTHVHDWSQGAVRLLTDTTDWPATGAPRRAAVSSFGISGTNAHTIIEQAPEPQPEDAA
FT      TAQDDAAGSTPATAPVVPVLLSGRTPDALRGQAAALRAALDTGRRPDLLDLAHS
FT      ATTRAGFEHRAVLLATDHPALTDGLTALADADDPAAPAWITGTTTRAETRLAVLFTGQG
FT      AQLGAGRELAARFPAFATALDAALDAFTPHLDRPLREVLWGTDAALLDRTAYAQPALF
FT      AVEVALYRLIESFGVRPDHLAGHSVGEIVAAHLAGVLSLADAATLVAARGRLMQALPDG
FT      GAMIAVQASEADVAPLLAGHEDQVAIAAVNGPSAVVLSGAEATVTALAEQLAADGRKTR
FT      RLRVSHAFHSPLEMPMLDAFRAVVEDLTLQPPLLPVVSNLTKPATVAQLTSADYVVDH
FT      VRHAVRFADGIDWLARHDTTAFLELGPDGVLSSAMAQDCLDAADADAVTLPALRAGRPEE
FT      HTLT'TALAGLHVHGATLDWTGCFAGTGARRTDLPTYAFQRRRYWPKALQSGTADLRSVG
FT      LGAAHHPLLSAAVSLADAGGTLTGRLSRQTHPWLADHTVRGTTLLPGTAFLELAVRAG
FT      DEVGCDRVEELTLAAPLLLPEQGGVQVQLWIGNPDVSGRRTVNVHARPDGTDDTPWTAH
FT      ATGVLTTADASRQLPASSEQGGTPLAGDHPALDAAQWPPAGAEPPLDGHYDRADGG
FT      FGYGPFVQGLRAAWRGDVVYAEVELPEAGRSDAEAFGLHPALLDAALHAAPTGLGER
FT      GRGGLPFSWEGVSLHAGGATTLRVRLTPVADDALALTVDAGTGAPVLSVDSLVRSVAT
FT      QQLDTAAAVARDALFRLDWTVPQPTATDPGPVALLGADPFGLLTHAGFADAPAYPDLAA
FT      LAAADGPVPTTVVLSLAGTGDDAADPARSAHRCAAEALAAVQTWLDHHERFAAARLVFV
FT      TRGATVGRDVA AAAAVWGLVRSQAQSENPGCFALVDLDPDGAVGAAALVAALVSGEPQLAV
FT      RGDVLRVARLVRRPLTEVGAGADGTGDGVDGSGVSFSGEGAVLVTGGTGGLGAVLARH
FT      LVAEYGVDRDLLLVSRGERAVGAGELVAELAGVGARVRVACDVTDRAAVVELVGGHAV
FT      SAVVHAAGVLDDGMVGALTGERLSAVLRPKVDVAVHLHEATRGLDLDAFVVFSSLAGVF
FT      GSPGQANYAAANAFDALMTRRRRAEGLPGLSLAWGPWEQSGGMTGTLTVDVAERLARS
FT      VPPLSVAQGLALFDAAVAGTATCVPVRLDLPVLRARGEVPPLRLSLIRVARRAAVAG
FT      SATAGNLAQRLRLRLDEDGRDEMVLDLVRGQVALVLGHATGGDVDAGRAFRDLGFDLSLA
FT      VELRNRLNTVTGLRLPATLVFDYPTVRHLATYVLDELLGTDAEVATVQPAAVAVADDPI
FT      VIVGMACRYPGGVSSPEDLWRVLTETGDAVSGFPTNRGWDVESLYHPDPDHPGTSYTRS
FT      GGFLHEAGEFDPGFFGMSPREALATDSQQRLLLESSWEAIERAGIDPVSLRGSRTGVFA
FT      GVMYSDYSAMLASPEFEGFQGGSSPSLASGRVAYTLGLEGPVAVTDTACSSSLVAMHW
FT      AMQALRSGEGLALAGGVTVMSTPAVFVDFARQRLSPDGRCKAFADAADGVGWSEGVG
FT      VLVLERQSDAVRNGHEILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALASGGLTAG

```

DVDVVEAHGTGTTLGDP IEAQALLATYGRDREPERPLLLGSVKSNLGHTQAAAAGVAGVI  
 KMLVAMRHGVVPRTLHVDAPSSHVDWSEGAVELLSEQAAPETGRVRRAGVSSFGISGT  
 NVHVIVEQAPGAKAIAAAGAARRTPGAVPVLLSGRGRSALRGQAARLLGHLQARPDAEL  
 VDVALSLATTRSFRFEQRAAVVAQDRDQLIASLGALAADRPDPAVVEGEAAGRGRRTAVLF  
 TGQGSQRAAMGRELHEVQPEFAAAFDCAVFDPLDRPLREVVFADGSDDEAALLDET  
 GWTQPALFAVEVALFRLVESWGVPRPDFVAGHSIGEIAAAHVAGVLTLEDACRLVAARAT  
 LMQALPTGGAMIAIQATEDEIAAHLDDTVAIAAVNGPQSVVISGDEEAAETIAATFAER  
 GRKTKRLRVSHAFHSPRMDGMLDAFRIVAEGLTYPRIPLVSDLTGRRADDAEVCTAE  
 YWVRHVREAVRFADCVRTLRLDAGATTFLLELGSDDLTTAMAEDTLGDDHDAELVPMRLRAG  
 RAEELAAATALARLQVRGVDVDWAAYLAGTGARRTDLPTYAFQHAYYWPQLPTPAAALA  
 AADPADQQLWAAVERGDARELADILGLGEQDLTPLDSLLPALTSWRRGNQEKHLLDTLR  
 YRVEWTRL SKPTAPVLDGTWLLVASDATAADQPALLDGLADALGSHGARVRRLLDDSC  
 ADRAVLAERLARTADVDAATQVLSVLPDERDADDCPPLTRGLALTVALVQALADTGAQ  
 GRLWTATRGAVSTNPADPVTHPVQAAAWGLGRGVALEHPRLWGGDLVLPQVFDERAGQR  
 LAGILAVKDAPDGEDQVALRATGVSGRRLVRHTVEALPTAAEFTATGTVLITGGTGGLG  
 AEVARWRLARAGAQHVLVTSRRGPDAPGAAELRAELEGYGPSVSVVACDVADRDALAAVL  
 TALPEELPLTGVVHTAGVGHYGPLDTLSTAEFAGLTAAKLAGAAHLDAALLADRELDFFV  
 LFGSIAGVWGSNGSAYGAANAYLDALALHRRARGLAATSVAWGPWAEAGMAADDAVSE  
 TLRRQGLGLLD PAPAMTELRRAVVRQDVTVTADVDWQRYAPLFTSARPSAL IAGLPEV  
 RALAADERTEQDATGASEVVTRVRLAEPEQLRLLTDLVRTE SATVLGHSSADAVPEGR  
 AFRDVGFDLSLTAVELRKRLGAATGLSLPSTMVFDYPTPLELAQYLRAEILGAVLEVAGP  
 VATGGADDEPIAII GMACRFPGGVSSPEQLWDLVASGTD AISEFPVNRGWQTGHLFDPD  
 PDRPGTTYSTQGGFLHEADEFDPTFFGISPREALVMDPQQRLLLETTWESFERAGIRPE  
 TLRSTLTGTFTVGSSYQEYGLGAGDGTGHEMVTGSSPSVLSGRLSYVFGLEGPAVTVDTA  
 CSSLVALHLACQSLRNGESNLAVAGGATIMTTPNPFIAFSRQRALAKDGRCKAFSDDA  
 DGMTLAEGVGVVLRVERLSDAQRNGHPVLAVLRGSAINQDGASNGLTAPNGPSQQRVIRQ  
 ALANARLAPGDIDALEAHGTGTPLGDP IEAQALFATYGRDRDPESALLLSVKSNIIGHT  
 QSAAGIASVIKMMALRHSELPPTLHADAPSSHVDWSAGTVRLLTQARAWPETGRPRRA  
 AVSSFGISGTNAHVLLLEQAPVADTPAEERPAVAPVPIAAGVVPWVVTARSAAALRGQAE  
 RLLAHAETVGTALPAAGPLDIGLSLVSARARFEHRAVVVPPAGTDPLAALRAVATDGPS  
 PVVARGVADVEGRTVFVFPQGSQWVGMSQLLDES AVFAERIAECAALAEFTDWSLV  
 DVLRGVVGAPSLERVDVQPASFAVMVSLAALWRSRGVLPDAVVGHSGQEIAAAVVS GA  
 LSLRDGARVVALRSQAIGRALAGRGGMMSVALSVDVLEPRLVEFEGRVSAAVNGPRSV  
 VVAGEPEALDALHARLTADDIRARRIAVDYASHSHQVEDLHEELLEVLAEAPRTSEVP  
 FFSTVTGDWLDTARMDAGYWFNRNLGRVRFADAVADLLAAEYRAFVEVSSHPLVLSMAVQ  
 EAIDEAGVPAVAAGTLRRDQGGTDRFLLSAAEVFVRGVDVDWAGLFEGTGASRIDLPTY  
 AFQHEHLWAVPPAPEAVAAADPDAAFWTAVEDGDVSALTAALGTDEDSVAAVLPALTS  
 WRRARRDRSTVDWRYRVAVWKPLGGTLPHPSLTGTWLLVTADGIDDTDVAGALETYGAE  
 VRRLVLDEECVDRAVLRERLAGAEDVTGIVSVLAAAERTDAVPGTSLVLGTALTVALIQ  
 ALGDAEIDAPVWALTARGAVSTGRADELTA PVQAQVTGIGWTAAL EHPQRWGGTLDLPAA  
 LDARAAQRLAAVLSGALGSDDQLAIRPSGVFTTRIVRAEATAGRPAGTWT PRGTTLVTG  
 GSGTLAPHLARWLAQGA EHLVLI SRRGTAAPGAAELVAELAESGTEATVAACDITDRD  
 AVAALLADLKADGRTVTRTVHTAATIELHTLDATTLADFDRVLHAKVTGAQVLAELLDD  
 EELDDFVLYSSSTAGMWGSGAHAAYVAGNAYLAALAEHRRANGLPALSLSWGIWADDLKL  
 GRVDPQMIRRSGLFMDPQLALSGLQRLDDNENVLAVADVWETVHPVYTSGRPTPLF  
 DEVPEVRRLTAAAEQSAGTVAEGEFAAALRALSDAEQQRTLLETVRTEAASVLGLSSAE  
 DLTDQRAFRDVGFDLSLTAVGLRNRLASVTGLTLPSTMVFDYPNPAALAYLHGELAGAR  
 SAAAGAAAVPTGAPDADDPIAIVGMSCRYPGGVGSAEDLWRIALDEVDAISGFPADRGW  
 DAEGLYDPDPDRPGRTYSVQGGFLRDVAEFDPGFFGISPREALSMDPQQRLLLETAWEA  
 FEHAGIDPVGQQRSGRTGTFTVGASYQDYASGVNSEGSEGHMITGTLSVLSGRVSYLFG  
 FEGPAVTLDTACSSSLVAMHLACQSLRNGESSLALAGGVSMSTPMSFVGF SRQRALAE  
 DGRCKAYADGADGMTLAEGVGLVLLERLSDARANGHVLA VIRGSAVNQDGASNGLTAP  
 NGPSQQRVIRQALANSAPVPGDIDVLEGHGTGTALGDP IEAQALLATYQGDRAPERPLL  
 LGSVKSNIIGHTQMASGVASVIKLVRLAQEGVVPKSLHIDRPSTHVDWSSGAIGLLTERT  
 PWPETGRPRRAAVSSFGISGTNVHTILEQAPADEAPTADPPRDGLVPVLLSGRGEAAL  
 RAQAARLLAFVEERPEAHLTDLAHSLATSRAALERRAAVIAADRDTLTRGLRALS DGRP  
 DPGLVQGTAGRGRTAFLFTGQGSQRPGMGRELH DRYPVFADALDEVLARLDDGPDRLR  
 EVLFAAPDSAEAAALLDRGTGYAQPALFAVEVALFRLLT SWGLTPDYLAGHSVGE LAAAHV  
 AGVLSLDDACTLVAAARGRLMQALPEGGAMVALEAAEDEVLP LLEGLTDRVSVA AVNGPR  
 SVVVAGVEEDVLLADLFAADGRRTKRLRVSHAFHSP LMDAMLDDFAAVARGLTYHPPT

FT IPFVSNVSGGLATAEQVRTPDYWVGHVRAAVRFADGIDWLATQGDVHTFLELGPDPVLS  
 FT AMARESLTDPSTALLPTLRGDRPEEPALVTAVAAHAHGARVDWSGYFADHGARRTTL  
 FT PTYAFQRRERYWPDTTAATSAHTPGSALDAEFWAAVERDDVAALAASLDLDDATVTAMVP  
 FT ALTAWRRRRRGEQTELDWRYRVWTKPRGGATAPAALTGRWLVLVPHDQDRQDDATAAW  
 FT AADVETALGTTTTRVRLTVTTTDRALAARI TEAAGDQGPFSGVLSLLPLATGDAGHPGAP  
 FT AALTLTTTAVQALGDAGIDAPLWNVTRGAVAVGRAEQVTAPEQAAVWGLGRAVALELPA  
 FT RFGGTLDDLPLATLDGQAARRLRAVLAATDGEDAVALRPSGVFLRRLAHAPAGPDTARTAF  
 FT DPAAGTVLITGGTGGIGGHVARRLARDGATHLLTSRRGPAAPGADALRAELEELGARV  
 FT TLAACDAADRDLAALLAELPDDAPLCAVFHTAGVVEDHVVDALT PENFAAVLRAKTVA  
 FT AHHLHEL TADLDLAAFVLFSSSTAGVLGAAGQGNYYAANAHLDALAEHRRSHGLTALSVA  
 FT WGPWAGSGMVADAAELTDRVRRGGFEPLAPEPAVRALLRAIENDDTTVALADIDWERFQ  
 FT RAFAAVRPLPFVADLPETGRATPATATGAATGLRQQLAELPEHERPAAVLDLLRTQVAA  
 FT VLGHADPRTVEDDHAFRDLGFDLSLTILELRNALNAATGLSLPATLVYDLPTPREMADFL  
 FT LAELLGTLPTD TAATVASTASPKLSASFEQGGTPFDDPIAVIGICRFPGGVTTPEELW  
 FT QLLDEGRDGISRFPDDRWDLAALGAGASDTLEGGFLTGVADFDARFFGISPREALAMD  
 FT PQQRLLLETTWEALERAGIDPTTLRGSTTGTVFVGTNGQDYPTLLRRSASDVAGYVATGN  
 FT TASVMSGRLSYALGLEGPAVTIDTACSSSLVALHWAGRALRAGECDLVVAGGVSVMASP  
 FT DSFVEFSTQGG LAPDGRCKAFSDAADGTAWSEGVGILVLERLSAARRNGHQVLGLIRGT  
 FT AVNQDGASNGLTAPNGLSQQRVIAQALADARLRPADIDAIEAHGTGTTLGDP IEARALI  
 FT TAYGRDRDAERPLLLGTVKSNIIGHTQAAAGAAGVIKMLMAMRHGTLPRTLHVGTSPSSHV  
 FT DWSGGTVALLDARWPRTGQPRRAGVSAFGVSGTNAHVVEQAPETEAPAAPAAEPAP  
 FT EATPTVVPWVVSGRSREALQAQLDRLTAHTAAHPARSAADVGRSLATDRTLFPHRAVLL  
 FT AGPDGVREAAARAAAPRTPGRTAFLFSGGQAQHALMGHDLYQRFVYADALD TVLAQFDT  
 FT VLDVPLRAALFAAPGTPEAALLDQTGFTQPALFAVEVALFR LAESWRILTDFVAGHSIG  
 FT EIAAAHVAGVFSLEDACTLVAAARASLMQQLPRDGAMVALEATEDEVAPLLTDGVALAAV  
 FT NGPRSVVAGAEDAVRAVADR LAADGRTRRLTVSHAFHSP LMDPMLTDFARVAEGLTY  
 FT HEPRIPLVSTLLGAPAGAE LRTPDYWVRHVRETVRFADGVRALHDAGAGTFVEIGPDGV  
 FT LTALTQQTLD TVEAGAPAVVPLQRRDRAGDLALLEGLATLHTHGTGPSWPAYFEATGG  
 FT HRTDLPTYAFQRRERYWPELGAPVATAPQDPAAWRYHETWAPLPAPAEAAAPAGRALVLVP  
 FT AGNRDTAWMTADEALGADTVTAEPDALAEQLTAAGDTPWRVVSLLAAASEGLPADGA  
 FT WPAALLATLDEAGVHAPLWCVTRGAVAVAGEAPTAVGQAALWGLGRVAALDHPDRFGGL  
 FT ADLPADTDAHAAGLLAAHLAAPGTEAEI AVRATGVHARRLVRTPAAADGATWLPTGTVL  
 FT VVGGTGGTGMTGGRAARWLVRGARHLVLTAPDGTTTAADTEALTAELAALGARITVVD  
 FT HDPTAPDGF AALLDGLPDDTPLTAVVYAPEADAAPGTAAELSAALAPVTALGAALTGRP  
 FT LDAFVLFGS IAGLWGVGRAAEAASGAYLDAFARACDRGTPALAVANGAWADLVGPSL  
 FT AAHLRMNGLPVMADTALTALSRAVADGSAAEAVADVWRWETFAPLHHEARRTALFDALP  
 FT EARGALAEAAADRADRKTAAGDYGRWLAEQPAADHDAILLALVTEKAATVLGHADHDLL  
 FT EPDLPPFDLGFDSLTAVDLRNQLTAETGLTLPATLVFDHPNPAALAAHLRAQLLGEASD  
 FT SAAPVAAPVALGADDDAIVIVGMACRYPGGVTS PEDLWQLVGDEVDVAGDFPTDRGWDL  
 FT AALAGDGPGRSATAQGGFLYDATDFDPGLFGISPREALVMDPQQRILLETSWEALERAG  
 FT IDPATLRGSGTTGVFVGGSGDYRPEEAGQWQTAQSASLLSGRLAYTFGIQGPTVSVD  
 FT TACSSSLVALHLAAQALRAGECSIALAGGVTVMATPVGFVEFSAQGALS PDGRCRAFSD  
 FT DANGTGWSEGVGMLVVERLSDARRNGHRVLAVLRGSAINQDGASNGLTAPSGPAQQRVI  
 FT RQALANARLRPADIDAVEAHGTGTRLGDP IEAQALLATYQQDRERPVLGLSKSNIGHT  
 FT QAASGVGGVIKMLVAMQHGE LPRSLYAENPSSHVDWTAGRAHLLTARTPWPDSSGRPRRA  
 FT AVSSFGASGTNAHAILEQPPREELPARPADDGAPLPFLLSGRSQNALRAQARRLLARLT  
 FT AHPDTRAADLAYS LATTRA AFHRAAITATDHDGLRTGLTAVAEGTTAPHTAEHHLQGT  
 FT GKRAVLFSGQGSQRLGMGRELHERHPVFAEAFDSVLARLDDRLDTPLRDVVWGTDEEAL  
 FT HATGNTQPALFAVEVALYRLIESWGVRPDFVAGHSVGE LAAHVAGVLSLDDACRLVAA  
 FT RAALMQRLPAGGAMIAVEATEDEVTPLLTGDVSLAAVNGPTAVVLSGAGDAVTALGQAL  
 FT AERGHRTTTLRVSHAFHSHLMDPMLADFR TVAEGLEYHPPRI PVVSNLTGDVADAADLC  
 FT SADYWVRHVRTVRFADGVRTMADRGVHLFLELGPDAVLSAMARQCAPDAVVVPALRRN  
 FT RDEDET LVGAVARLHVHGAGPRWDAYFAGRGAQWLDLPTYPPQRGRFWPESLPGAASAA  
 FT PAAGQPAETDAAFWDAVAQEDFTALESVLDVESDALS KVL PALMDWRSRQADESQLAGW  
 FT RHRIVWKRLTGAALAHRKALSGTWLAVVPEGFADDPWVTTTTLDGLGTHLVHLEVAEADR  
 FT AALADAIAARTADGTRFGGVISLLALREELTGAVPEGTALTTTTLLQALGDAGVDAPLWC  
 FT VTRSAVSAGRTDRPHRPLQGA VWGLGRVAALEYPQRWGGLVDLP EEPDERSAAGLAAVL  
 FT AGLDGEDQVAVRGTA VLARRLVPAPGRKPSRPWHPSGTVLVTGGTGALGAHVARRLAKD  
 FT GAQHLVLLSRRGPDAPGAELRAELDALGTDTVAAACDVADRDQLTAVLDALPADRPLT  
 FT GVVHTAGVLDDGVLDRLTPERFQEVFRAKVTSALLLDELTRDRELA AFVLFSSASA AVG

```

FT      NPGQANYAAANAVLDALAEQRRVLGLPATSVSWGAWGGGGMADADGADEAARRAGVGAM
FT      DPHLAVEALLRLVAEKEPTAVVAEVALDRFAGAFGGSRPSALLREFPGYREALAAQAEQ
FT      AADGGGLAARLAALPPARRLDTVVDLVRTRAAQVLGYPDTEAVAAERSFRDLGVDSLGA
FT      VELRNQLSAATGLNLPATLVFDHPTPLVLGEHILGGLFPDEPAGSDDETEIRALLASVP
FT      LDQLREIGVLEPLLQLAGRGGRAADGDDGESVDSMTVADLVRAALNGQSDL"
FT      CDS      34384..50691
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4X2"
FT      /note="polyketide synthase"
FT      /transl_table=11
FT      /gene="nysJ"
FT      /function="responsible for condensation steps 15 to 17 in
FT      the nystatin polyketide backbone synthesis"
FT      /product="NysJ"
FT      /protein_id="AAF71767.1"
FT      /translation="MNAPENPETPENNVAALRAAVKETDRLRRQNRMLVAAAKEPIAV
FT      VGMACRFPGAVDSPEALWEMVATGTDVISGFPDDRWDLEALRNSGTDARDTDVSQRGG
FT      FLDCIADFDPGFFGISPREAVTMDPQQRLLLLTAWEAVERAGIDATTLRATRTGAFIGT
FT      NGQDYAYLLVRSLLDDATGDVGTGIAASAASGRLSYTLGLEGPALTVDTCSSSLVALHL
FT      AVQALRNGECEGMALAGGVNVMATPGSLVEFSRQGLARDGRCKAFADAADGTGWSEAG
FT      VLLLERLSDAQRNCHPVLAVVRGSAVNQDGASNGFTAPNGPSQQRVIRQALANAGLATG
FT      DIDAVEAHGTGTPLGDPIEAQSILATYQDRAHPVLLGSIKSNMGHTQAASGVAGVIKM
FT      IMAMRHGVLPRTLHVDRPSTHVDWTTGSELLTDAHPWPETGRPRRTGISSFGVSGTNA
FT      HVIVEQAPDTPAEAAADTPPRTPLPWLARSARTGAALRDQATALLDHLDRPDGDRGPT
FT      ALDTAFSLATTRAALAEHRLAVVTGTDGTAGRDALTAWLAHGTAPDAHEGHAAGRTRCAA
FT      LFSGQGAQRLGMGRELHARFPVFARALDTAVDLLDAELGGTLREVIWGTDDAPLNETGF
FT      TQPALFAVEVALYRLIESWGVAPDFVAGHSIGEIAAAHVAGVFSLEDAC TLVAARAGLM
FT      QALPRGGAMVAVEATEDEVSPLLTDGVAIAAINGPTSLVVSGDETATLAVARLAEQGR
FT      RTTLRVSHAFHSPMLDPLMAEFRAVAEGLSYGEPQIPVVSNLTGAVADGTLTGADYW
FT      VRHVREAVRFADGIRALTDAGVGAFLELGPDGTLAALAAQQSAPDAVSVPLRKORDEEP
FT      AAVAALARLHTAGVPVDWTAIFYAGTGAHRTDLPTYAFQYERYWPKATYRPADATGLGLT
FT      AADHPLLGAAMSVAGSDELLLTGTLSLATHPWLADHVVGGMVFFPGTGFLAVRAADQ
FT      VGCDRVEELMLAAPLILPATGTVMQIAVGAADDDGGRDLRFFTRPGDDPDAAWAQHAT
FT      GRITEGERVLALDTTTWPPRDAEPVDIDGLYDRYRANGLDYGPVFRGLRAVWRDTEIY
FT      AEVALPEGTADADAFGLHPALFDAVLHSTLFASADGDDRSLLPFAWNGVSLHAAGADAL
FT      RVRITSCGPDAVEITA VDPQGRPVVSVESLTLRAAGPDAGTADHRADAGSLFRMDWTPR
FT      TVHAPATPATWAVLGTDPIGLTEALTAAGPDTVTGLRDGVDALGELTAGDDRVPVDPVVA
FT      VPLRGATDHGPAGAHDLTRTVLALLQEWLAEEFARSRLLLVTRGAVADGERGPLDLAA
FT      APVWGLVRSQAQSENPGRLLLVDLDDTAESAAQLPLLPALLDADEPQAVVREGTVRVGRL
FT      ARLDSGRGLVPPPGTPWRLGSRAGSLDGLALLPHPEARPLTGHEVRVGIRAAGLNFR
FT      DVLNALGMPYPGDAGLFGSEAAGVVVEVGPEVTGLAPGDRVMGMLFGGFGPLGIADARLL
FT      TPVPADWSWETGASVPLVFLTAYYALKELGGLRAGEKVLVHAGAGGVGMAAIQIARHVG
FT      AEVFATASEGKWDVLRSLGVADDHIASSRTLDFEAAFAEVAGDRGLDVVLNALSGEFVD
FT      ASMRLLDGGRGFLEMGKTDIRAADSVPDGLSYHSFDLGMVDPEHIQRMLLDLVELFDRG
FT      ALAALPVRSDVRRAGEAFRMSLAQHIGKIVLTVPQPLDPDGTVLLTGGTGGLAGLLA
FT      RHLVTEHGARHLLLAGRRGPDAPGAAALHAELTALGAETVAACDVADR TALAAALLATV
FT      PAEHPLTAVVHTAGVLDGTLTALNPDRLATVLRPKVDAAWHLHDLTRHLDLAAFLVYS
FT      STAGVMGGPGQANYAAGNTFLDALAAHRHALGLPATSLAWGAEQAGMTGALT DHDLR
FT      RVSDAGGQPLLTAERGLALYDAATAADEPLIVPLGLTGGLPAGVGVPVAVLRGLVRTAG
FT      RRARAGTAGVSRAGLAERLAALPEEERTPFLVELVRTEATVLGHGSTDPVDARREFRQ
FT      LGFDSLTAIELNRNLGKATGLTLPATLIFDYPTPDR LAVHLHDELLGADAPVTVTAAAQ
FT      AADPEHDPVVI VGMSCRFPGGVSSPEELWDLVASGTDAITGFPADRAWDRHPQLAGAPG
FT      ARTGQGGFLRDIADFDAAFFGISPREALAMPQQRILLEVAAEAAERAGIDPQTLRGSD
FT      TGVFMGVSGQDYAGLVMSRDDIAGHATTGLAVSVVSGRLAYALGLEGPALSVDTCSS
FT      SLVSLHLAAQALRAGECTMALAGGVTVMTTAAANFTGFSRMGGLAQDGRCKAFSDSADGT
FT      GWSEGA AVLVLERLS DARRAGHRVLAVVRGSAVNQDGASNGLTAPNGPAQQRVIRQALA
FT      NAGLTPVDVDAVEAHGTGTPLGDPIEAQALIAAYGTD RDPEHPLLLGSVKSNIIGHTQSA
FT      AGAAGLVKMVMAMRHGILPQTLHLTEPSSHVDWSAGTVRLLTERTAWPRTDRPRRAGVS
FT      SFGISGTNAHVILEQPPAEPTPAADPGRPAPTVVAWPVSAQTPAALDAQLDRLRTAAAL
FT      APLDTAHTLATGRSLFEHRAVLLATVGD PATGAPDLPEVARGAATPHRTAFLFSGQGAQ

```

```

FT      RSGMGRELHAAFPVFAAAFDEVVAVLDAELGSDADGGVSLREVMWGGGSELLDRTRFTQ
FT      PALFAVEVALFRLVASWGVGPEFVAGHSVGEIAAAHVAGVFSLVDACRLVVARASLMDA
FT      LPVGGVMVAVEAAAEVVP LLVDGVAIAAVNGPVSVVVSGVEAAVGQVVDQLVERGRRV
FT      RRLAVSHAFHSPLMDPMLDAFRAVAEGLEYHQPRI PVVSNVTGEVAAAEEELCAADYWVR
FT      HVRATVRFADGVRTLAERGATAFLEIGPDGVL SALARGVLP AEALVTPTLRKDRDEESA
FT      LLAGLARLHVAGVTV DWSAALTGTGARGTDLPTYAFQ RERYWP ELAAEPAGGGADAADA
FT      EFWAAVERADAT ALAAHLDIDG DQLGAVLPALSAWRTRRR TTSATNALRHRESWEPLSL
FT      AGTPHTGGV LVLVPA AATTDPWVADVVAALGPDARRVDVPADGTDR AALAALLTEAADD
FT      TAPTAVVSLALDE TSGDDAVPAGTTATAALVQALADTGAPAPLWALTRGAVAALPDEQ
FT      PTAPAQA AVWGLGRIAALELPRHWGGLVDLPADLDERTARRLPAALADAGDEDQLALRA
FT      TGAYGRRITPAPAPDDAPGTGWQPTGTVLITGGTGALGRHTARWLA AHGAEHLLLSRS
FT      GPDAPGAAELTTEL TALGARVTLVACDAADREQLTRVLA EVPRDCPLTGVVHTAGVLDD
FT      GVLTLGLTPDRFATVFRAKVASAVLLDELTRD TDLAVFALFSSVAGAVGNPGQAGYAAAN
FT      AVLDALAARRRAQGLAGTSIAWGAWAGDGMAARHTRPGAEPVGLLDPDLAVPALARAVT
FT      EPQPTLV LADLQQPRLLESLLALRPSPLLSRLPAARTAA RAVQEADRRRAGAAADLRDQ
FT      LAGTAPADRHA VLLRLVRTTAAAVLGHTGADAI RADKPF RD LGFDSL TAVELSSALAAA
FT      TGLALPPSLVFDHPSPRALADHLRAELTGDRPESAPAAPPAPVPAADDDPIVVVGMACR
FT      FPGGVTTPEEFWQLLAEGRDGIDAFPTDRGWDLDVLGRRRPGPQRPPRSAASSYDAAAF
FT      DPGFFDISPREALAMPQQRLLLETAWEAVERTGTDPTRLRGSRTGVFVGTNGQDYAGL
FT      VLRAQEDVEGHAGTGLAASVISGRLAYAFGFEGPAVTVD TACSSSLVALHWAVQALRAG
FT      ECSLALAGGVTVMTTSTSFAGFTRQGGLAPDGHCKAFSDSADGTGWSEGVGLVVERRS
FT      DALRNGHEILAVVRGSAVNQDGASNGLTAPNGPAQQRVIRQALANAGLAPGDVDAVEAH
FT      GTGTVLGDPIEAQALLATYQDRPADRPLWLGSVKSNI GHTQAAAGAAGLMKMLALQH
FT      GTLPRTLHVTEPSTRVDWSAGAVRLLTERTVWPRTDRPRRAGVSSFGISGTNAHVILEQ
FT      GPAEPTPTAPADRPTRTPAVL PWVVSARSATALDAQ LARLRAFAAERPDLPPADVAHSL
FT      VTSRATFEHRAVLLAAPDGITAAARA EARERSTAF LFSGQGAQRSGMGRELHAAFPVFA
FT      AAFDEVVAVLDAELATGSGGGVSLREVMWGGGSELLDRTRFTQ PALFAVEVALFRLVAS
FT      WGVGPEFVAGHSVGEIAAAHVAGVFSLV DACRLVVARASLMDALPVGGVMVAVEAAEAE
FT      VVPLLVDGVAIAAVNGPVSVVVSGVEAAVGQVVDQLVERGRRV RRLAVSHAFHSPLMDP
FT      MLDAFRAVAEGLEYHQPRI PVVSNVTGEVAAAEEELCAADYWVRHVRATVRFADGVRTLA
FT      ERGATAFLEIGPDGVL SALAAACLFDTDAEVVPALRKGRPEEHTAL TAAQLHVAGVDI
FT      DWTAVLAGTGRRIALPTYAFQ RERYWPSLAAQAPGDAGGLGLEAGRHP LLGAATTVAG
FT      SAEILLTGRLSTTAQPWLAVYEADGRTVLPAAVLAELAVRAGDQADCPTVAELTVAAPL
FT      VLTGAAAQRLQVRVAAPDDTGRRALSVHARPDDSPDSPWTLHATAVLTHTDPQPPAPDT
FT      GWPPERAVPLDALPTATGPARIAAAWQWGDELCAEIELPEPGPAERAFALHPALLDTAV
FT      RAGGLLDG DATL DALGWRGLALHAASATALRVRLTPDGTDTWALEATDPQ GAPVVSVTG
FT      LTLGTPTVDRSGAGAADDGATLLDLEWVPAPQAAPTGGDHL PYAVLGDQLAELDGQLRI
FT      AGDGPGRVASLAALLDGGAPLPRLVLAPVLGVPTGEGDLPAAVRGTTTAVLELLQRWTA
FT      DARTADSHLVI VTRGAVAAGAEDVHD LAAAPVWGLVRS AQSEHPGSFLLLDLPADPAG
FT      ASRAAAPATL AALLDAGETQAAVRADTLTVARLTRAADGPEATAGHPVRDWRDGTVLI
FT      TGGTGGLGGLLARHLVTGHG IKHLLLAGRRGPDAPGARALRDELAALGA ETVAAACDVA
FT      DRAALDRLLAQLPPEHPLTAVVHTAGVLDDATVGT LTPERLDTVLRAKADA AWHLHDAT
FT      RDRDLAGFVLYSSVAGVTGGPGQGNYAAGNTFLDALAAHRAAQGLPGLSLAWGPWGQDA
FT      GMTGT LGAADLARLERSGMPPLTPEQGLALFDAAGARGDGF AVAVRLARGAAAPGADEV
FT      PAVLRALVRGRRRTAAAAGHAGVLARRLAALDAEQRHQALLDLVRTETA AVLGHSGADA
FT      VPAERDFNRLGFDSLMAVELRTRLATATGARLPATLVFDHPTPDAVARHLASTLPGGTA
FT      AGPDRSPLAELDR IAAELSP EGADDATRQGVVGRRLRHLLAQWDGTRQDGGGTTVDDRIE
FT      AASAEVLAFIDHELGRQADS"
FT      CDS      50747..56947
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4X1"
FT      /note="polyketide synthase/thioesterase; contains a
FT      C-terminal thioesterase domain"
FT      /transl_table=11
FT      /gene="nysK"
FT      /function="responsible for the last condensation step in
FT      the nystatin polyketide backbone synthesis"
FT      /product="NysK"
FT      /protein_id="AAF71768.1"
FT      /translation="MPDEKKLV DYLKWVTKDLHQTRQRLQEVEAGRHEPV AIVGMACRF

```

```

FT      PGGVRSPEDLWELLSAGRDGIGPFPADRGWDLAALAGDGPGRSATQEGGFLLPDAAAFDP
FT      GFFDISPREALAMDPPQRLLLLETAWEAVERSGIDPAGLRGSRTGVFVGTNGQDYAHLVL
FT      AAQDDMGGYAGNGLAASVLSGRLAFALGLEGPAVTLDTACSSSLVTLHLAAQAVRAGEC
FT      GLALAGGVTVMTTSSSFAGFSLQGGGLAPDGRCKAFAEAADGTGWSEIGIGLLLVERLSDA
FT      QRNGHPVLAVLRGSVAVNQDGASNGLSAPNGPSQQRVIRQALAGAGLVPGDVEDAVEAHT
FT      GTRLGDPIEAGALLATYQDRPADRPLWLGSVKSNLGHTQAAAGVAGVIKMLALRHGV
FT      LPQTLHVDAPSSHVDWESGAVRLLTAPVAWSEGDDRVRRAGVSSFGISGTNAHVILEQA
FT      PDQPEPTAEETAAAPGGTAEEERAAAPVAPRAVPWPVAARTAGALDAQLVVRALTTAP
FT      GRTAADVGHALATARTPFHRALLVHEGGAVTEVARGAVPTGDRGGLAVLFSGQGSQRP
FT      GMGRELHARYPVFAAAFDETVALLDARLGTSLRDIVWDQDRTRLDDTRHTQPALFAVEV
FT      ALYRLLAŠWGIRPDHVTGHSIGEITAAHVAGVLTADACTLVAARATAMSELPFGAMV
FT      ALEATEDEVRLPLLTDDLAIAAVNAPRSVVVAGAEDAALAVRRHFDDLGRRTTRLPVSHA
FT      FHSPLMDPMLDAFRTALAPLTFAEPEIPVVSNLGLPATAEELATPHYWVCHVRQAVRF
FT      GDGVRALADRGVRTFLELGPDGVL SALVRENLEPEGLVAVPVLRKERPEETTVLAALGT
FT      LWAHGADVWDVAVFAGTRTPQADPVELPTYAFQRARYWPTLGARHGDPADLGQTAAAHF
FT      LLGAAVTLADADETVLTGRLALPSHPWLGDHRS DGRITVPGVAFELAVRAGDLSGTPH
FT      LARLDLPAPLTLGDGDTVTLQVRVGAPDPAGHRPLTVHARLAATEDAPWTTTCATGLLAP
FT      DAPEAPADPIGPADAGWPPERDARVPVADLDAAATAAGRHYGPHFQGLTGLWRRDGEVF
FT      AEVALPTATAADRAFGIHPALLATALRATAALDDDHDTAGHTPEPTGITGLALHATGATA
FT      LRVRLTATGPDTVLAAAADATGGAVLTADTVTLGSPQDRPAPAPAGHTGQGGFLHLDVW
FT      PVDPGSRATGTRWAVVGDDDELGLYALHRADETVSAYAAASLGGAIGDSGLAPDVFLVPV
FT      VGGPDAGPDVAVHAVTARALGLLQEWLNEPRLAGARLVFVTRGAVAVPGETVTDPAAGAV
FT      WGLLRSAQTENPGSLLLVDLDDAFRSAGMLPHVLTLDLDEQQLVVRDHAVRAARLARLEP
FT      AAGTAPARAWDPDGTVLITGGTGGLGAALARHLVTVRGARHLLLAGRRGPEAPGAGELV
FT      AELTAQGADVRVAACDVGDRTALDALLATVPAAHPLTAVVHTAGVLDDALIGSLTPDQL
FT      ATVLRPKADAAWHLHDATRGLDLAGFVLYSSVSGVLGSPGQGNYYAANAYLDALARHRA
FT      DQGLPALSLAWGPWGRGSGMTASVSDADLERMARGGLPPLTVEDGLALFDAAVGRPEPA
FT      LVPSRINVAGLRDQALPALWRDLVPRARRTAATADRS PVTVRERLRHLDETGQEQLLI
FT      DLVVGYTAGLLGHPDPTAVDPERGFLELGFDSLVSVGLRNQLAEILGLRLPSSIVFDSK
FT      SPVKLARWLHQELANGPQPGATGPAAADARPAVRSDDTLEGLFYNAVRRGKLEAMRML
FT      KAVANTRPMFDTPAELEELSEPVTADGPGRPRLI FVSAPGATGGVHQYARIAAHFRGS
FT      RHVSALPLMGFAPGELLPATSEAAARIVAESVLMASEGEPFVMVGHSTGGS LAYLAAGV
FT      LEDTWDVRPEAVVLLDTASIRYNPGE GNDLDRTRFYLAIDIDSPSVTLNSARMSAMAHW
FT      FMAMTDIQAPAPTPTLLVRAARALDGRFLDTSSVPADEV RDIDADHLSLAKEHSALTA
FT      QAIEGWLAELPDPA "
FT      CDS      57095..58279
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4X0"
FT      /note="P450 monooxygenase"
FT      /transl_table=11
FT      /gene="nysL"
FT      /function="presumably involved in modification of the
FT      nystatin macrolactone ring"
FT      /product="NysL"
FT      /protein_id="AAF71769.1"
FT      /translation="MSTPTAPPSLKA E V P P V L R L S P L L R E L Q S R A P V C K V R T P A G D E G W
FT      L V T R H T E L K Q L L H D D R L A R A H A D P A N A P R Y V H N P F L D L L V D D F D L A R T L H A E M R S L F T
FT      P Q F S A R R V M D L T P R V E A L A E G V L A H F V A Q G P P A D L H N D F S L P F S L S V L C A L I G V P A E E Q
FT      G K L I A A L T K L G E L D D P A R V Q E G Q D E L F G L L S G L A R R K R I T P E D D V I S R L C L K V P S D E R I
FT      G P I A S G L L F A G L D S V A S H I D L G T V L F I Q H P D Q L A A A L A D E K L M R G A V E E I L R S A K A G G S
FT      V L P R Y A T A D V P I G D V T I R A G D L V L L D F T L V N F D R T V F D E P E L F D I R R A P N P H L T F G H G M
FT      W H C I G A P L A R V N L R T A Y T L L F T R L P G L R L V R P V E E L R V L S G Q L S A G L T E L P V T W "
FT      CDS      complement(58378..58572)
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4W9"
FT      /note="ferredoxin"
FT      /transl_table=11
FT      /gene="nysM"
FT      /function="participates in electron transfer in P450
FT      monooxygenase systems"

```

```

FT      /product="NysM"
FT      /protein_id="AAF71770.1"
FT      /translation="MRITVDPGRCVAGQCVLTA PDLFDQDDDGLVTVLAGAADAADPG
FT      DVRDAAALCPSGAISVAAD"
FT      CDS      complement(58637..59833)
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4W8"
FT      /note="P450 monooxygenase"
FT      /transl_table=11
FT      /gene="nysN"
FT      /function="presumably involved in modification of the
FT      nystatin macrolactone ring"
FT      /product="NysN"
FT      /protein_id="AAF71771.1"
FT      /translation="MSTEADARTAAPQCPVAFPLRRPGRPFPPPEYATYRGGAGLVRSE
FT      LPSPGVVLVTRHEDVRAVLTDPRISADPSKPGFPKAGRTGGAPSQYEVPGWVAMDPPE
FT      HGRFRKTLIPEFTVRKVRELRPVIQQIVDERIDAMLAAGTSADLVESFALPVPSLVISS
FT      LLGVPKVDRDFFEDRTRVLVRLSSTDEERDKATQALLRYLGRLIQIKQRRPGDDLISRL
FT      IAAGTLRQELSGVAMLLLIAGHETTANNIGLVVQLLTNPRWIGDDRIVEELLRYYSV
FT      ADLVAFRVAVEDVEIGGQLIRAGEGIVPLIAAANHDATAFAAPSEFDPERSARSHVAFG
FT      YGVHQCLGQNLVREEMDIAYRTLFA RIPSLTLAVPVEELPLKYDGVLFGLHELPVTWK"
FT      CDS      complement(59830..60888)
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4W7"
FT      /note="putative aminotransferase"
FT      /transl_table=11
FT      /gene="nysDII"
FT      /function="presumably involved in mycosamine biosynthesis"
FT      /product="NysDII"
FT      /protein_id="AAF71772.1"
FT      /translation="MSFTYPVSMPWLQGRELDYVTEAVGGGWISSQGPYVRRFEEAFAA
FT      YNDVPFGVACSSGTTALTALRALGVGPGDEVIVPEFTMIASAWAVTYTGATPVFVDCG
FT      DDLNIDVSRIEEKITPRTKVIMPVHIYGRQCDMDAVLNLAYEYNLRVVEDSAEAGVVRP
FT      RGDIACFSLFANKIISAGEGGVCLTHDPLHAEQMAHLRAMAFTKDHSFLHKKLAYNFRM
FT      TNMQAAVALAQTEQLDTILALRRDIEKRYDEALRDI PGITLMPPRDVLWMYDLRAERRD
FT      ELCAYLAGEGIETRVFFKPMRQPGYFSADWPALNAARLSADGFYLPHTHTGLTAQE QEF
FT      ITGRIRAFYGVA"
FT      CDS      complement(60909..62429)
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4W6"
FT      /note="putative glycosyltransferase"
FT      /transl_table=11
FT      /gene="nysDI"
FT      /function="presumably responsible for attachment of the
FT      dGDP-mycosamine to the nystatin macrolactone ring"
FT      /product="NysDI"
FT      /protein_id="AAF71773.1"
FT      /translation="MTLPSGNTRLGWRRRRMHSPGDRAGRVRGARARRPATFRGVLSMG
FT      ANRRPILFVSYAESGLLNPLLVLAGELSRRDVADLWFATDEKARDEVAADV DGS PVRFA
FT      SLGDTV SQMSAVTWDDATYAEVTQSRFRKAHA AVIRHSFAPESRMAKYRRL EEEIVEEVE
FT      PALMVI ESMCQFGYELAITKGIPFVLGVFPVPSNVLTSHVPFAKSYTPSGFPVPHSGLP
FT      AAMSLAQRIENQLFRLRTLGMFLTSDVRKVVEEDNRVRTELGIAPQARQMMARIDHAEQ
FT      VLCYSVRELDYPFPMHPKLRLVGT MVPPLPQAPDDDGLSDWLSAQKSVVYMGFGTITRL
FT      TREQVASLVEVARRLDGRGHQVLWKLP RQGQELLPPAAELPDNLRIEGWVPSQLDVLAH
FT      PNVKAFFTHAGNGYHEGLYFGKPLVVRPLWVDCDDQAIRGQDFGVSLTLDRPETVDTE
FT      DVLDKITRVLDQPSFTERAEHFAGLLRDAGGRAAAADLLLGLPALATD"
FT      CDS      62659..66759
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4W5"
FT      /note="polyketide synthase; contains KS domain with

```



```

FT      Cys->Ser substitution in the active site"
FT      /transl_table=11
FT      /gene="nysA"
FT      /function="presumed loading module of the nystatin PKS
FT      complex"
FT      /product="NysA"
FT      /protein_id="AAF71774.1"
FT      /translation="MTIGADEDPVVVVGMACRYPGGVAGPEDLWELVRTGRDATTAFPD
FT      DRGWDLAALAGDGPGRSATREGGFLTGAADFDAFFGMSPREAVSTDPQORLVLETAW
FT      ALERAGIDPHSLRGSRTGVFVGASGQDYAAVTHASPDLDGHALTGLAPGVASGRLAYV
FT      LGLEGPVAVTVDTTSSSSLVALHWAVRALRAGECSTALAGGVTVMSTPAAFVGHTRQGGL
FT      APDGRCKPFSDDADGTAWAEGVGIVVLEHLSTARAAGNPVLAVLRGSANVDGASDGLT
FT      APSGPAQERVIRAALADARLAPADIDLVEAHGTGTRLGDPVEARALLAAYGQDRDPDRP
FT      LRLGSLKSTLGHAAQAAGIGGVIKTVLTLRHGLMPRIHRLATPTRQVDWSQGAVAPLTD
FT      HTPWPPADRP RRAGVSSFGISGTNAHVILEEAPPADVPTVRPGTLRPSTVPWPVSAATP
FT      EALDAQLARLRAHLRTHSDLDPLDVGYSLATGRAALRHRAVLLPPADGTAADAVEHARG
FT      AAHQRRTAVLFSGQGSQRPGRMRELAARFPVFADALDDALRALDRHLDGPPREVMWGT
FT      AALLDRTGWTQPALFAVEVALHRLVASLGVTPDFVGGHVSVEIAAAHVAGVLSLEDACR
FT      LVAARATLMQALPAGGAMAALAEATEDEVAPLLGAHLALAAVNGPTAVVVAGAEDAVRQL
FT      TARFADRGRRTSRLAVSHAFHSPLMEPMLDAFRDVVSRLTFHQPSIPLVSNLTGELAGS
FT      EITSAEYVWRHVRDTRVFADGITALAKAGADVLIELGPGGVLSAMARDTLGPDSTTDV
FT      PALSCKRPEETAFAAGALGRLHTLGVPVDWPAFYAGTGARRVELPTYAFQHVHRHWPTPPR
FT      PNGAGPGALGHPLLGSAVELADGGGTVCSGALSLRTHPWLADHTVAGRVVLPATALLEL
FT      AVRAGDEAGCDVLHELHLTTPPALPDAAALHVQVHVGPADTTGRRAVTVHTRPDHHPAG
FT      DWTRCATGTLGSTPPSAEAATGGTPAAWPPADAEPDLADHYERLADRGFDYGPTRFRG
FT      LRAAWRRGAEIFADVECPPGTADDAPDHGLHPALLDAARHAAMAVDGTVPVAWHGVRLH
FT      AVGATALRVIRIPTTTGTTLTLTAVDVHGAAPVVTVEALTARPLTDEERAAPRTPRQARGE
FT      TPADARPARPAAARPGPAGEPLPDDTTGSHPTAGHLAALPPAARERQLLDLVRTQAAAVL
FT      GHGPPEAVGTRSVFKELGFDLSLAGVELADRLTARTGLRLPATLVFNFTPERAAHRLGE
FT      LLAATAPLDPGAYGEELTRFEAIVTNLPQDGPERRAVADRLDAIVSALRQNSPAEVPSS
FT      DEDIDTVSVDRLLDIIDEFEFTT"
FT      CDS      66805..76383
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4W4"
FT      /note="polyketide synthase"
FT      /transl_table=11
FT      /gene="nysB"
FT      /function="responsible for condensation steps 1 and 2 in
FT      the nystatin polyketide backbone synthesis"
FT      /product="NysB"
FT      /protein_id="AAF71775.1"
FT      /translation="MQEPQQGQPDQQEKIVDYLRRTSDLRARRRRIGELESKDNEPIA
FT      IVGMGCRLPGGVNSPESLWDLVRSGGDAISGFPVDRGWDLETLTGNGDGSSATHEGGFL
FT      YDAAEFDAAFFGISPREATAMDPQQRLLLEVAWEALERAGIAPTALRGSRSGVFGVSYH
FT      WGAPSADAATELHGHALTGTAASVLSGRLAYTLGLEGPVAVTVDTACSSSLVALHLAAQS
FT      LRVGESSLAVIGGVITLTPSVFVEFSAQGGLAPDGRCKAFSDAADGTGWAEGVGVLVA
FT      ERLSDAQRNGHPVLAVLRGSANVDGASNGLTAPNGPSQERVIQQALARTGLTPADIDA
FT      VEAHGTGTRLGDPIDIAQALLATYQGHTPDQPLWLGLSKNIGHTQAAAGVAGVIKMVM
FT      ALRHGHLPPTLHADAPSSHVDWSAGSVRLLTEGQWPETGRPRRAAVSSFGISGTNAHA
FT      LLEQAPHPADTADAGDDAAPTEPAGAPALPWIVSGHSPQALRDQAAALAAVETDPAL
FT      RPQDIGHTLHTARALLERRAVVAPDRAELLAATHELAAGRSANAVVEGLADVEGRTVF
FT      VFPQGSQWVGMAQQLLDES AVFAERIAECAALAEFTDWSLVDVLRGVVGAPSLERVD
FT      VVQPASFAVMVSLAALWGSRGVLPDAVVGHSGQEIAAAVVSAGLSLRD GARVVALRSQA
FT      IGRALAGRGGMMSVALSVDVLEPRLVEFEGRVSVAAVNGPRSVVAGEPEALDALHARL
FT      TADDIRARRIAVDYASHSHQVEDLHEELLEVLAEAPRTSEVPFFSTVTGDWLD TARMD
FT      AGYWFRNLGRVRFPADAVADLLAAEYRAFVEVSSHVLTMAVLDLIEEAGVTAVATGTL
FT      RRDQGGAGRFLLSAAEVFVRGVDVDWAGAFEGTGAARVDLPTYAFQRRERYWNTRTAADR
FT      TPADAPMDAEFWAAVEQADVSAALTAALGTDEDSVAAIPLGLTSWRRARSQRTTLD SWRY
FT      RVTWTPLAQVPRATLTGTWLLVTTDGDIDTDVAGALESYGAEVRLVLDEECTDRAVLR
FT      ERLAGAEDVTGIVSVLAAAEDDAARHPGLTRGLALTVSLVQALGDAEATAPLWFLTRGA

```



```

FT      FATGSPDPVTRPLQSQIAGVGWTTALEHPQRWGGTVDLDPDLDARAAQRLAAALSGALG
FT      AEDQLAVRAAGVLARRIVRAGHRAGRPARTWAPRGTTLITGGSGTLAPQLARWLAERGA
FT      EHVVLVSRRGADAPGAPELIAEAAESGTEVTVAAACDITDRDAVAALLADLTADGRTLRT
FT      VIHAAAAIELSALADTTVAEFADVHVHAKVTGARILDELDDDAELDDFVLYSSTAGMWGS
FT      GVHAAVYVAGNAYLSALAEQRRRARGRLRTTSHWGWKPPDDRARELADPHRIRRSGLLEYLDP
FT      ELALTALQHVLDDETVIGLMDIDWDTYHDVFTAGRPAHLFDQIPEVRRRLDQASVPDP
FT      AGPAADGLAARLHGLAAAEQDRLLLTLVRTEAAAVLGHASAESFPERRAFRDLGFDSTV
FT      AVDLNRNLVAGTGRLRPLSTMVFDHPNCAALAAFLKTTALGVPGAAPQQAATGTPADDD
FT      PIAVIGMSCRYPGGAATPEELLRLALDGADVISEFPADRGWDARGLYDPDRPGHTYS
FT      VQGGFLHEAAGFDPGFFGISPREAVAMPQQRLLLETSWEAFERAGIDPASLRGSAAGT
FT      FFGASYQDYSSTVQNGTGESEAHMVTGTAASVLSGRVSYLLGLEGPVTVDTADGMTLA
FT      ALHLACQSLRDGESSLALAGGAAMATPHAFVGFSSRORALAKDGRCKPFSDTADGMTLA
FT      EGVGVVLLERLSHARANGHRVLAVIRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAG
FT      LTGADVDAVEAHGTGTLGDPIDIAQALLATYQDRDAERPLLLGSVKSNIIGHTQAAAGV
FT      AGVIKMLAMDAGELPGTLHLDPSSSHVDWTAGAVELLRGRTPWPESGRPRRAGVSSFG
FT      ISGTNAHLILEQAPATEPPADPDRLRDTATDTPVPWPLAAKSPAALRAQAARLLATVEH
FT      DPDFPAPVGHALATTRAALHRAVVVGERREDFLRGLAALSTGASTAGLVSGIAGPDP
FT      EGAVFVFPQGGSQWGMGRELLATSEVFRTAIDDCATALAPYVDWSLHDLVLAGEGDPAL
FT      LERVDVVPQALFAMMVGLSALWRSHGVVPAAVVGHSQGEIAAACVAGALSLADAARVVA
FT      LRSQALPQLSGRGGMMSVSAPVERVTALLAPWQEALSVAAVNGPSSVVVSGDTDALDAL
FT      HTACQEQGVRAKVSVDYASHGRHVEAVRDELARVLAPVDPRAPEVFPFYSTVTGDRVDD
FT      AAFDGAYWYTNLRQTVRMEEATRALLAAGHRVFIIEVSPHPVLAAPIQETQEAVAEATGG
FT      SAVVLGSLRRDEGGPRRFLTSLAEATHGAPVDWTTTFARSAYQPVLDPTYFPQRQDFW
FT      PEARPATPAAGADASDAFWQLVENQDLAALADALGVPADDEHTALGTVLPALSAWRK
FT      AQARTRIDELRYHVQWTRVAEPAAAPTGRLLVAVPPDHADAPWVAALDALGTDTVRF
FT      EAKGTRACGWAQAQIQLVEDGEEFTGVVSLAAAEDLHPDFGSVPLGLGQTLVLVQALG
FT      DAGLTAPLWCLTRGAVATGRDDALDSPTQGALWGLGRVVALEHPDRWGGLIDLPLATLDA
FT      RAAARLTGLLADPAGEDQLAVRATGVLARRMVHAAPSAPRTGRRWRGRGTCLITGGTGG
FT      IGGRVARWMAEHGAHLVLTSSRRGPADPGAAALRAELEALGARVTLAACDVADRDLALAA
FT      LLADLPADQPLTSVFHSAGVADGDARAADLTLDQLDALLRAKLTAHHHLHETAPLDLD
FT      AFVLFSSGAAVWVGSGGQPGYAAANAYLDALAAHRRSLDLPASVAVGTWGEVGMATVPE
FT      VHERLHRQGVRAPEPDHAIGALQQMLEDDDTTAVTLMDWAEAFAPSFTATRPSALFSTV
FT      PEAVRAVTGDPGTAGDDVDSATPPLRRHLEELSAAERGRALVEAVRAEASATLGHDTP
FT      DAIPAGRAFRDVGFDSTAVELRNRLRTALGLPLPALVFDHPTPTALAGHLGALLFGT
FT      APEDAGTGRPDDPDARIREALATVPIGRRLRKAGLLDMVLKLADGATDAPAPEADAPSE
FT      SLDDMDAEALLRLATENSAN"
FT      CDS      76403..109693
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4W3"
FT      /note="polyketide synthase"
FT      /transl_table=11
FT      /gene="nysC"
FT      /function="responsible for condensation steps 3 to 8 in the
FT      nystatin polyketide backbone synthesis"
FT      /product="NysC"
FT      /protein_id="AAF71776.1"
FT      /translation="MSTNPDKYVEALRSSLKEIERLRRQNEQLVAAAVEPVAVVGIGCR
FT      FPGGVTSPEDLWELVAEGRDVIGPFPQDRGWDLEKLAGGGEGGSLAQVGGFVEDAAGFD
FT      PGFFGISPREAVAMPQQRILLEITWEALERAGIDPSTLRGTPTGVFVGTGQDYGEVI
FT      KASAEDVEVYSTTGHAAVISGRLSYTLGAEGPAVTVDTCSSSLVALHWAVQALRGGE
FT      CSMALAGGASIMATPGPFVAFTAQSGLAADGRCKPFSDRADGTGWGEGAGMLVLMRLSD
FT      AQREGRPVLAVLRGSAINQDGASNGLTAPNGPSQQRVIRAALDSAHLTAADIDAVEAHG
FT      TGTTLGDPIDIAQALLATYQDRPRPLWLGSVKSNIIGHTQAASGAAGVIKIMIMALQRGVL
FT      PRSLHATEPTTVDWDTAGSVDLLDETVAWPETGRARRAGVSFGISGTNAHVILEQAPT
FT      APEEPTTEPTVRPAVVPWALSARTAAALDAQRARLTGHLADTPDADPLDVGYALADGRA
FT      TFEHRAVLLPDGTELAHGTAGEGPCAVLFSGQGSQRPGMGRELHARFPVFAAAFEITA
FT      LLDTHLDRPLREVWGTADLLNDTGWAQPALFAVEVALYRLVASLGVTPDFVGGHSIG
FT      ELAAAHVAGVLSLEDAC TLVAARARLMQALPRGGAMLAI RATEDEVTPH LTDDVSI AAV
FT      NGPTS VV VAGTEEAVAAIGARFTAQDRKTTRLRVSHAFHSP LMDPMLAEFRAVAAGLTY
FT      HEPRI PVL SNLTGTVA AVADLCSADYWRHVREAVRFADGVTALTDRGVTTLVELGPDG

```

VLSAMAQESLPDGA AAVPLLRKDRPEELSAVTGLARAHVVRGVTVRWAGLFDGTGARRAD  
 LPTYPFQHQRFWPTAARAAQDVTAAGLGAADHPLL GATVELADGAGYLF TSRLSVRTHP  
 WLADHGVQGRALLPGTAFVELAVRAGDEAGCDRVEELTLAAPLVLPERGGVQLQVRVGA  
 PDAAGRRTLGIFSRVEDGFDLPWSQHATGVLTAGAGAPDPTFDATVWPPSGAEPVDLTG  
 AYERLAALGFQYGP AFQGLRAAWRRDTEVYAEVALPDGADTDPAAFGLHPALLDAAQHA  
 AAYADLGAISRGGLPFAWEGVSLAAAGATTVRARIAPAGEDTVTI AVYDAAGGTVLSVD  
 SLVSREVPADAPGAAGTVHRDSL FHVWTP LQGRPGPAPATVAVLGPDPDALADTLRAT  
 GIRTTAPRDLAALADAEGVPDLVVTLLTTTPGAPVPDAAHATTA AVLALAQOWLADDR  
 FADARLVLVTRGATDGTDPAAAAAGGLIRTARTENPGRFALLDLAPDTGRPD PETLATA  
 LAASHDEPD LAVRGTDVHAARLARVPLATEPTTWNPDGTVLITGGTGGLGAVLARHLVA  
 THGVRHLLLASRRGPAADGADDLTAELTGLGATVHIAACDVADPAALADLLGTV PAGHP  
 LTVVVHTAGVDDGVLGSLTPQRLD TVLRPKADA AAWHLHEATRHLDDLDAFVLFSSVAAT  
 LGSPGQANYAAGNAFLDALAARRAATGLPATSLAWGPWTQSVGMTSSLSLDVERIARS  
 GMPPLTLEQGTALFDAALAAGPAALAPVRLDLPVLR TQGDIA PLLRGLIRTPVRRTAAQ  
 VSQTADGLAQRLAGLDAAARREALLELVRTQIAQVLGHADATEVETGRQFQDLGFDSL T  
 AVELRNALNTATGLRLPATMVFDYPTPHALADHLRDELLGTEAESTTAVVPVTRTAGTD  
 DPIVIVGMACRYPGGIASPEDLWRLVSQGADATGPFPTNRGWDDLNL YDPDPDRPGRTH  
 VRAGGFLHDAGSFDADFFGMSPREMATDSQQRLLLELSWEAVERAGIDPASLRDSGTG  
 VFAGVMYNDYGTTLTGDEYEAFRGN GSAPSVASGRVSYTLGLEGP AVTVDTACSSSLVA  
 LHWAAQALRAGECSLALAGGVTVMSTPSTFVEFSRQRLAPDGRSKAFEAADGVAWSE  
 GVGMLVLERQSDAVRNGHEILAVVRGS AVNQDGASNGLTAPNGPSQQRVIRQALASGGL  
 STADVDAVEAHGTGTTLGDP IEAQALLATYGRDRDPENPLLLGS IKS NIGHTQAAAGVA  
 GVIKMVMAMRHGVL PQTLHVDAPSSHVDWSVGAVELLTEQTAWPETGRARRAGVSSFGI  
 SGTNAHVVIEQSP TAVPATPASADRSVEEPPAVPWALSGKTPDALRQDQAARLLAHVEAH  
 PALRPVDISYSLIATRTAFDHRVAVLGTDRAEALRAL TALAAGETDPAALTGTVTRGRT  
 AFLFSGQGSQRLGMGRVLYERFPAFAEALD TVLTALDAELGHPLRDI IWGEDAQLVDRT  
 GYTQPALFAIEVALFRLLEAWGITPDFVAGHSIGEIAAAHVAGVLSLGDACRLVVARAV  
 LMQSLPEGGAMIAVQATEDEVLP LLTDDVSIAAVNSPTS VVVSGYENATLAVARHFADQ  
 GRRTTRLRVSHAFHSP LMAPMLDDFRAVVE SLTFTAPTTPVVSNL TGELAPAEALCSAD  
 YWVRHVREAVRFADGIRTLADRGVTTFVELGPDSVLSAMAQESAPEGAGTI PLLRRDRP  
 EEQAVLAALCHLQVLGVEADWSATFRGLDPVRVDLP TYAFQHRWFPAARPARPDDVRA  
 AGLGAAEHPLLGA AVQLPDDDGALFTGRLSLRTHPWLADHTVLGTVLLPGTALVELAVR  
 AGDETGS GHLEELTLAAPLTL PEDGATLLQVRVGSADDTGRRTVTVHARPDDTADRTWT  
 LHATGVLATTPPAAAAFDTTVWP PADAEPLTTDDCYAHFTTHRFAYGPAFQGLRAAWRA  
 GDVLYAEVALPESATDEAAAFGLHPALLDAGLHAALLADDRDTGLPFSWEGVTLHASGA  
 TALRVRLAPNGPNGLSVTAADPAGNPVATVTRLLARPLDAEQLT IHSALTRDALFHLDW  
 TPVPLPDTANSAPPALLGPDTAVLADALGDP AVARHATLDDL LAGDTPPATV LVPLGA  
 PLDGDTAQHAAH LTRSALT LVQQLATDRLADSR LVFVTHGAVATDDAPPTDLAAA V  
 GLIRSAQTENPGTFTLLDLDEPDSTTALS RALTLEDPQLLLRAGRARAARL TRTPAPT  
 TTTHTPWSADGTVLV TGGTGGLGGLVARHLVRSCGVRHLLLSRSRGVGAAGAAGLVAEL  
 ESLGARVVAAACDVGDGSAVAELVAGVSESYPLSAVVHAAGVLDDGVVGS LTPERLA AV  
 LRPKVDGAWN LHATRGDLDAFVVVSSVAGVFGGAGQANYAAGNAFLDALMVHRVAGG  
 LPGVSLAWGAWDQGVGMTAGLTERDVRRAESGMPLLTVDQGVALF DAALATGSAA LVP  
 VRLDLAALRTRGDIA PLLRGLVRAPLRRTAATGLATGADTGLVQRLGRLDHAQRHEALL  
 DMVRSSAALVLGHADGNAIDAERA FRDLGFD SLTAVELRNRLRTATGLHLSATMVFDHP  
 TLSALAEHLRDELFGAVESEVRVPVQALPPTADDP I VVVGMACRFPGGVTS PEDLWRLV  
 DDGTDAITTFPTNRGWDDLNL YDPDPEHFGTSYTRSGGFLHEAGEFDPAFFGMSPREAL  
 ATDSQQRLLLESSWEAIERAGIDPLTLRGSATGVFAGVMYSDYGS ILGGKEFEFGQGG  
 SAGSVASGRVSYALGFEGPAVTVD TACSSSLVALHWAAQALRAGECSLALAGGVTVMST  
 PSTFVEFSRQRLAPDGRSKAFEAADGVGWSEGVGILVLERQSDAVRNGHEILAVIRG  
 SAVNQDGASNGLTAPNGPSQQRVIRQALASGGLSTADVDAVEAHGTGTTLGDP IEAQAL  
 LATYGRDRDPENPLWLGLSKS NIGHTQAAAGVAGVIK MVMAMRHGVL PQTLHVDAPSSH  
 VDWSVGAVELLTEQTAWPETGRVRRAGVSSFGISGTNAHVIVEQPALVESPA AEPSGRE  
 PGVVPLPLSGKSPEALRDQAARLLAGLAERP ALRPLDLGYSLATTRS AFDHRAVVLATD  
 RADAVRAL TALAAADADLSAVVG DTRTGRHAVL FSGQGSQRLGMGRELYERFPVFAEAL  
 DVAIDHLDAALPAQASLREVMWGDDVELLDETGWTQPALFAVEVALFRLVESWGV RPDF  
 VAGHSIGEIAAAHVVG VFSLEDACRLVAARATLMQALPTGGAMIA IQAAEDEV TQHLTD  
 DVSIAAVNGPTS VVVSGAES AARTVADRLAENGRKTTRLRVSHAFHSP LMDPMLAEFRA  
 VAEGLSYATPTLPVVSNL TGRLATADDLCSAEYWARHVREAVRFADGVSTLENEGVTTF  
 LELGPDGVL SAMAQQLTGDAATVPALRKDRDEETSALTAL AHLHTAGLRVDWAAFFAG

FT SGATRVLDLPTYAFQHATYWPTGTLPTAHAAAVGLTAAEHPLNGSVELAE GEGVLFTGR  
 FT LSLQSHPLADHAVMGQVLLPGTALLELAFRAGDEAGCDRVEELTLAAPLVLPERGAVQ  
 FT TQVRVGVADDTGRRRTVTVHSRPEHATDVSWTQHATGTLTMSAPADTGFDATAWPPADA  
 FT EPLATDDCYARFTTLGFAYGPVFQGLRAAWRAGDVLVYAEVALAESTGDEATAFGLHPAL  
 FT LDAALHASLVAHEGEESNGGLPFSWEGATLYATGATALRVRLTPTGTGDGRSVAIAVADT  
 FT AGRPVAIDNLVSRRVSGDQLTGAAGLARDALFTLDWNVPENLVPEPVNPENTGGGHA  
 FT QDQDGRPAAATVALVGADGTAIAADLTAAGIHTTLHPDLTTLATTDADVPKTVLIPLTG  
 FT TGTGTGTGTESTDGIGTGAAESDASAPSPA EVAHTLSTAALALQEWTAQERFAGSRLA  
 FT FVTTGATAAGGTDVMDVAAA AVWGLVRS AQSEAPDTFVLIDRDPGPAGTHDRATAAAERG  
 FT QLLLRALHTDEPOLALRDGGVLAARLARFDTA AALTPPADRAWRLDSTAKGSLNGLALT  
 FT PYP AALAPLTGHEVRVEVRAAGLNFRDVLNALGMPGDDVGSFGSEAAGVVVEVGPEVT  
 FT GLAPGDQVMGMITGSFGSLAVDDARRLARLPEDWSWETGASVPLVFLTAYYALKELGGL  
 FT RAGEKVLVHAGAGGVGMAAIQIARHVGA EVFATASEGKWDVLRSLGVADDHIIASSRTLD  
 FT FEAAFAEVAGDRGLDVVLSLAGDFVDASMRLLGDGGRFLEMKTDIRAADSVDPGLSY  
 FT QSFDLAWVVPETIGTMLAELMDLFRGTALRPLPVRTWDVRHAKDAFRFMSMAKHIGKIV  
 FT LTLPRSWKPEGTVLVTGGTGGLGGLVARHLVRS CGVRHLLLT SRSGVGAAGAAGLVAEL  
 FT ESLGARVVVAACDVGDSAVAELVAGVSESYPLSAVVHAAGVLDDGVVGS LTPERLAAV  
 FT LRPKVDGAWN LHEATRGLDLDAFVVFSSVAGVFGGAGQANYAAGNAFLDALMVHRVAGG  
 FT LPGVSLAWGAWDQGVGMTAGLTERDVRRAAESGMPLLTVDQGVALFDAALATGSAALVP  
 FT VRLDLAALRTRGDIAPLLRGLVKAPIRRAAATTPGDTGLAEQLTRLQRAERRDTLLALV  
 FT RDQAAMVLGHTSGDGVDP SRAFRDLGFDLSLTAVELNRNIGAATGLRLPATAVFDYPTAD  
 FT ALAAHLLTELLGPDAESDPDEPGDPTAGPTDDPIVIGMSCRFPGDIGSPEDLWRL LGD  
 FT GADVVTDFPTNRGWDLDNLYDPDPAHAGTSYARTGGFLHDAADFADFFGMSPREAMAT  
 FT DSQQRLLLESSWEAIERAGIDPLTLRDSRTGVFAGVMYSGYGTRLDGA EFEGFGQGS  
 FT LSVASGRVSYTFGFE GPAMTVDTACSSSLVALHLAAQALRGECTLALAGGVTVMSIPD  
 FT TFIEFSRQRLAPDGRSKPFSEADGVGWSEGVGM LLLERQSDAVRNGHQILAVVRGSA  
 FT VNQDGASNGLTAPNGPSQQRVIRQALASGGLSTADVDAVEAHGTGTTLGDP IEAQALLA  
 FT TYGRDRDPENPLLLGSIKSNLGH TQAAAGVAGVIKMVMAMRHGVLPRSLNITEPSSHVD  
 FT WSAGAVELLTEQTAWPETGRARRAGISSFGISGTNAHVILEQPEARHSAPEEADTAEA  
 FT AAKAPATAHLVPMPWALSGKTPEALRAQAARLLAHLQORPELAPADIALSLATQRSQFT  
 FT HRAVVLSTDRDEATRAL SALATTAASDPSALTGTVTMGRCAVLFSGQGSQRLGMGRELY  
 FT ERFVPVFAEALDVVIDHLDAALPAQAGLREVMWGDDVELLNETGWTQPALFAIEVALFRL  
 FT VESWGV RPDFVAGHSIGEIAAAHVGVFSL EDA CR LVAARATLMQALPAGGAMIAVQAT  
 FT EDEVI PHLTDEVAIAAVNGPTSVVISGAE EATQTVAQHFADQGRRTTALRVSHAFHSPL  
 FT MMLAEFRVAEGLSYATPTLPVVSNLTGQVATADELCSAEYWRHVREAVRFADGV TAL  
 FT EAEGVRTFLELGPDGVLAAMARETVADDTVTVPVLRNMPEERTLLTALGR LHTTGTP I  
 FT DWAALLAPTGARPVLDLPTYAFQHRPFWPSGPRDTADAAAVGIAGASHPLNGI VELADE  
 FT EGLLFTGRLSLQSHPLADHAVMGQVLLPGTALLELALRAGDEVGCDHVEELTLAAPLV  
 FT LPERGAVQTQVRVGVADTTGRRRTVTIHSRPARATTTDS DTHGTDTPTWQTHATGVLVAG  
 FT LPATATVPFDATVWPPAHAE PVDLADFYASRAGEGFGYGP AFQGLRAAWRRDGEVFADV  
 FT ALPEAGRTEAEAYGLHPALLDAGLHAAWL VAPDGEPTRTGSPVFSWRGVFLAASGASSV  
 FT RVRLGRDSDGTYLSLAIDTTGAPVASVQALSMRTVSVTALSATAGLARDALFRLDWASA  
 FT PEPACQPDDTVTVI PAVAVVGTETSEL TSELTAALRAAGADV DVRTTSLTDEPAPALIA  
 FT LPLVASDQTGTAEAAPVPAAVHDLTRRALALVQTRLQE QHFADTKFVFVTRGATVGRDV  
 FT AAAAVWGLVRS AQSENPGCFALVDLPD GAVGAAALVAALVS GEPQLAVRGDVL RVARL  
 FT VRRPLTEVGAGADGTGDGVGGSGVSFSGEGAVLVTGGTGGLGAVLARHLVAEYGV RDL  
 FT LLVSRGERAVGAGELVAELAGVGARVRVACDVTDR AAVVELVGGHAVSAVVHAAGVL  
 FT DDGMVGALTGERLSAVLRPKVD AVWHLHEATRGLDLDAFVVFSSLAGVFGSPGQANYAA  
 FT ANAFLDALMTRRRRAEGLPGLSLAWGPWSLT DGTSGMLADAEADRLTRSGVPPLTAEQGL  
 FT ALFDAALATGDATCVPVRLDLSALRAQGEVPPLLRSLIRGRSRRAAAESATATGLRER  
 FT LVGLNPVERQEVL LLDLVRGQVALVLGHADADDVHPARAFRELGFDSLTSVELNRNLNTV  
 FT TGLRLPATMVFDYPTVEVLVSYVLDELLGTDAEVATVQPA AVAVADDPIVIVGMACRYP  
 FT GGVASPDLDLRLVTGVD AVSPFP TNRGWDVESLYHPDPDHLGTSYTRSGGLHEAGEF  
 FT DPGFFGMSPREALATDSQQRLLLESSWEAIERAGIDPVSLRGSRTGVFAGVMYS DYSAM  
 FT LASPEFEGFGQSGSSPSLASGRVAYTLGLEGP AVTVDTACSSSLVAMHWAMQALRS GEC  
 FT GLALAGGVTVMSTPAVFVDFARQRLSPDGRCKAFADAADGVGWSEGVGV LVLERQSDA  
 FT VRNGHEILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALASGGLTAGD VDVVEAHGT  
 FT GTTLGDP IEAQALLATYGRDREPERPLLLG SVKSNLGH TQAAAGVAGVIKMVLAMRHGV  
 FT VPRTLHVDAPSSHVDWSEGAVELLSEQA AWPETGRVRRAGVS SFGISGTNAHVILERPE  
 FT AARRPVMETNTVEPSTVPVWLSGKTPEALRAQA AKLLSSI ERPELRLVDVGMSLV TGR

```

FT      STFEHRAVVLAADRADAARALSAIAADEADAAAATGRVAGGRHAVLFSGQGAQRLGMGR
FT      ELYERFPVFAEALDVVDHLDAALPAQAGLREVMWGDDAELLNETGWTQPALFAIEVAL
FT      FRLVESWGVPRPDFVAGHSIGEIAAAHVAGVFSLEDACRLVAARATLMQALPAGGAMIAV
FT      QATEDEVTPHLTDDVAIAAINGPNALVVSQVEDAAVEIGARFAAEGRRTRRLHVSHAFH
FT      SPLMDPMLAEFRVVAEGLSYAAPSLPVVSNLTGQVATADELCSAEYWVRHVREAVRFAD
FT      GVTALEAEGVRTFLELGPDPVLAAMAGASLTSSLAVPLLRKDRPEEPAALAALQHLI
FT      AGARVDWPVLFAGVGAGRVELPTYAFQRGWFWPVGVRVGVGGDVGAAGVLSAGHPLLGAA
FT      VELAAGAGVVLTGRLSLSSHGWLADHAVMGRVFPVGTALLEMMVRAGDEVGCGRVEELT
FT      LAAPLVLPERGGVRVQVAVDAPDAAGRRGVGVYSCPDGVGQAVWSQHAGVVLASGVADQ
FT      VGGFGDGGVWPPQGAHSVDAEGCYELFADAGFGYGPVFQGLRAVWRRGEELFAEVALSD
FT      EVAESADTATGFGHPALLDASLHASLLSSLEGQSADGGPALPFAWEGVSLFASGATAL
FT      RVR LAPAGEHAVSVTAVDPTGAPVISIDALRTRRLTLDEVNASHTQLSDALFGVQWTTV
FT      PSTPAADHPSVAIIGTDHLGLAEALSSSSAGATTTTTTAAAYESLDALIAAGPEVSVPDV
FT      TLIGLTTEDAIAQYVNDHDATVAGQGTIGAGAAVDAARRLTAEALRTIQAWLADERLA
FT      ARRLVFTVTRGAADGQDVAANAQGLVRSQAQNTENPGTFGLLDLDGTEASTAVLGEALTSD
FT      EPQLLLRDGHLHAARLTRLASPADTAVPTEWNADGTVLITGGTGGLGAQFARHLVDYRG
FT      VRNLLLVSRGPDAPGTTELVAELTAHGAEVAVQACDVADGDAVAALVAGVPDEHPLRA
FT      VVHTAGVLDDGVIGSLTEERLATVLRPKADAAWHLHEATRGLDLDAFVVFSSVAGVFGG
FT      AGQANYAAANAFDLALMAQRRAGLPGLSLAWGPWDQGTGGMGLSDAEADRLARSGIP
FT      PLSAEQGLALFADALALAGTSTPDRAAGSAAASTSGTGDTIAIPAAALVAPVRLDLAAL
FT      AAQGEVPAIILRGLVTRTRRTAAGGSVTVAGLVNRLSGLTADERRQELLELVRTQAAV
FT      LGHADPASVDSTAQFRDLGFDLSLTAVELNRNLSTATGLRLTATLVFDYPNTDALAEHLR
FT      DELFGAVESEVRVPVQALPPTADDPVVVGMACRFPGGVTSPELWRLVDAGTDAITTF
FT      PTNRGWDLESYDPDPAHLGTSYTRSGGFLHEAGEFDPAFFGMSPREALATDSQQRLLL
FT      ESSWEAIERAGIDPLTLRGSATGVFAGVMYSYDGSILGKKEFEGFQGGQSAGSVASGRV
FT      SYTLGFEQPAVTVDACSSSLVALHLAAQALRAGECTLALAGGVTVMSTPGTFVEFSRQ
FT      RGLAPDGRSKAFEAADGVGWSEGVGILVLERQSDAVRNGHEILAVIRGSAVNQDGASN
FT      GLTAPNGPSQQRVIRQALASGGLSTADVDAVEAHGTGTTLGDP IEAQALLATYGRDRDP
FT      ENPLLLGSIKSNLGHTQAAAGVAGVIKMMAMRHGVLPQTLHVDAPSSSHVDWSVGAVEL
FT      LTEQTVWPETGRVRRAGVSSFGISGTNAHVILEQPEAVQRLAPGAAETVEPVAIKPSAE
FT      PSLVPWALSGKSPEALRAQAARLDRFLAERPEPRSIDIGHSLAVTRSQFDHRAIVLDD
FT      AKAPADSLAALASGVADPAVSDAVSTGGSVAVLFTGQGAQRLGMGRELYGRFPVFA
FT      EALDVVDHLDAALPAQAGLREVMWGDDVELLNETGWTQPALFAVEVALFRLVERWGV
FT      PDFVAGHSIGEIAAAHVAGVFSLEDACRLVAARATLMQALPTGGAMIAVQATEDEVTPH
FT      LTDEVAIAAVNGPSSVVISGAEEATQTVAQHFADQGRRTALRVSHAFHSPLMDPMLAE
FT      FRAVAEGLSYATPSLPVVSNTLGLWATADELCSAEYWVRHVREAVRFADGITTLEAEGV
FT      RTFLELGPDPGILSALAQQSLAGEAVTVPLRKRGEESTALTARAHLHTRGLIEDWQDF
FT      FAGVGAGRVELPTYAFQRGWFWPVGVRVGVGGDVGAAGVLSAGHPLLGAAVELAAGAGVV
FT      LTGRLSLSSHGWLADHAVMGRVFPVGTALLEMMVRAGDEVGCGRVEELTLAAPLVLP
FT      GGRVQVAVDAPDAAGRRGVGVYSCPDGVGQAVWSQHAGVVLASGAADQVGGFGDGGVW
FT      PPQGAHSVDAEGCYELFADAGFGYGPVFQGLRAVWRRGEELFAEVALSDEVAESADTAT
FT      GFGHPALLDASLHASLLSSLEGQSADGGPALPFAWEGVSLFASGATALRVLAPAGEH
FT      AVSVTAVDPTGAPVISIDALRTRRLTLDEVNASHTQLSDALFGVQWTTVPSTPAADHPS
FT      VAIIGTDPFGLADGLSDALPLVEERGLAALAASEHPVDPDLVLPVAGTRRTGVPADAE
FT      GHTDAGTSDMLRSVREATAQVLEQIQWLADDRFEAARLVFVTRGAVSVGEGGIADLAA
FT      SAVWGLVRSQAQSENPGCFGLLDLDLALDSDLAPEVDIERDRDRDPVGGTVQPALAAA
FT      LHATADEPQLALRGTVQAARLTRIAPQTDRAETDPAETDRPEIDTRRPGTVLITGGT
FT      GGLGGLLARHLVAERGVRSLVLASRGLAAEGAELVADLEALGAVVAVQTCADVADGDA
FT      VAALVAGVSDEYPLTAVVHTAGVLDDGVIGSLTEERLATVLRPKADAAWHLHEATRDL
FT      LDAFVVFSSLAGVGGAGQANYAAANTFLDALMAQRRAGLPVSLAWGPWDRAAGMTG
FT      TLSDAEADRLARSGVPPISAEQGLALYDAATAGERPLVVPVRLDLAALRGLGDPALLR
FT      GLVRTPARRTAAAGAAPSADVLTRQLAGLGGAEEVLLRLVRGQAAVVLGHADGSAIG
FT      AGRQFQELGFDSLTAVEFRNLNAATGLRLPATLLFDYPTPADVVGHRLRGLGTGEVSG
FT      AGSVLAALDNLEAVIAGLSLDDAGEHQLVAGRLEVLRKAWDMRSAEAVDGGADVDIE
FT      EASDDDMFALLDDELGLN"
FT      CDS      110113..110868
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4W2"
FT      /note="putative thioesterase"
FT      /transl_table=11

```

```

FT      /gene="nysE"
FT      /product="NysE"
FT      /protein_id="AAF71777.1"
FT      /translation="MTTSTEEESLWARCFHPAPAAPVRLFCFPHAGGSASFYFPVSAQLS
FT      SVAEVFAIQYPGRQDRRKEAGVSDLATLADQVYDALRPLLKERPSTFFGHSMGATLAFE
FT      VARRFEADDGDLVRLFASGRRAPSRVREEAVHRRSDDGIVEELKLLAGTNTALLGDEEI
FT      LRMILPAIRSDYQAIETYRCPPDVTVRAPLTVLTGDRDPKTSLDEAEAWRGHTTGDFDL
FT      KVLPGGHFFVSSEAPAIIDLLRAHLAGNG"
FT      CDS      111258..114158
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4W1"
FT      /note="transcriptional regulator"
FT      /transl_table=11
FT      /gene="nysRI"
FT      /function="transcriptional activator for the nystatin
FT      biosynthesis genes"
FT      /product="NysRI"
FT      /protein_id="AAF71778.1"
FT      /translation="MRKQSGSSGLLTTLVGRDDELRTLARHAAAARDGRAGLVLLHGPA
FT      GMGKTSLLRSFTASDVCRGMTVLVYGTCTGETVAGAGYGGVRELLGGLSGGDARRSPLL
FT      EGLAARALPALTADPAGPDAATGAYPVLHGLYWLAAARLMAQRPLVLVDDVHWCDESL
FT      AWIDFLLRRAEDLPLLVLAWRSEAEPVAPAVLADIAAQRRTPTVLGLHPLGPDDIGEMV
FT      RRVFRTTAAPSFVSRVAAVSGGNPLALARLLDELRAEGVRPDAAGERRAAEVGSHVLAR
FT      SVRCLLERRPPVWRGVARAIAVLGPECTELLAALAGVPAATVDEALLVLRAGILAADR
FT      VDFVHDVVRSAVLDDVAPPTLAELRTNAALLSDAGRPSEELAGQLMLLPVLDQPWMAA
FT      VLRDAAAQAESRGAPEAGVRCLYRVLEVEPDNVAVRIQMARALAEINPPEAMRLLKEAL
FT      SLAGDVRTRAQVAVQYGFTCLAVQESPSGVRMLEDALAELTAELGPEPGPVDRELRTLV
FT      ESVLLIVGADEKVTIGAVRDRAARLTMPPGDTPAQRQMLAMTTVLTAMDGRDARSVDQ
FT      ARRALRAPGVELEPWSLLSASFALSLADEVADAQYALDMLQYQGDNAAVWTYVLALST
FT      RALLHHGVGAFPEALADAQTAVEILGEERWADGAVLPRVALATALVDRGEPERAEHVLD
FT      GITRPRLERFVIEYHWYLQARAYARWVRGDFQGALDLLLACGRSLEESRFSNPAFVPPWW
FT      ADGAVLLATLDRHDQARELAAYGSELAERWGTARGLGAFMAQGVAAAPGRAGIDHLTEA
FT      VSLLDADSPARAMEARAELLGLGHAHLKRDDLRAAREHLRAAADLAQRCGAVKLGVDARKL
FT      LVTAGGRVRRMTASPLDMLTGMERTVADLAVTGASNRAIAEALFVTVRTIETHLTSVYR
FT      KLGVGGRAELSAVLETRTATSGRQPPAWVSQARGRA"
FT      CDS      114182..117043
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4W0"
FT      /note="putative transcriptional regulator"
FT      /transl_table=11
FT      /gene="nysRII"
FT      /function="presumable transcriptional activator involved in
FT      regulation of nystatin biosynthesis"
FT      /product="NysRII"
FT      /protein_id="AAF71779.1"
FT      /translation="MPRSKARNQPTTCTPQCAPDAHGDPTMLLECGREQRLIGDLLHRL
FT      GQGRPSVLSLTGRPGHAQNALVRWGACRARHDGLRVLRAQATPAERELRYGAVLQLLAV
FT      LDGPHGSTLDAAIRHDGPPPLPVPGIEEVLRRGTAPTLLVVEDVQWLDPASLTLWLQIL
FT      LRHLGPDTPLAVLASSCGDTTAFDTPKAPAVPGPPDTPVPVARFVVPALTDGRVAATVR
FT      AVCGTPGDEEFIAALTSATAGNPAILRDALRAFVDHGLPADADHLPHELHALTAGVVGDH
FT      TVRALDGLPAEVNAVLRALAVCGDLLDFHRVRALAGAHSLSEDRIRTLTLASVGLTVSVG
FT      DKVHIRFPASKARVIEDMPAAERADLYVRAAELTHSCGVNDEDVAHLLLRSSPLGAPWV
FT      VPLLRRGFAAALRREDHHRACACLSRALQEPLDPRERSLLTLELAAAEEAVARPEAGDRR
FT      LGELVRSTVADTDPTSSGEGVGVRaidLGFARGNSEWVRRTAGEALPYAGPADREELVA
FT      LFWLAAVRDDDAPMIPVVPRLPDRPVPPAQAGARAWQLATAGEDADKARKLARIALTGG
FT      VNESLMPKLAACAALFATDDNDEAVHGLDTMLTAARSAHLRSMAARIFNLRIHLCA
FT      ARLEAAERDLDSAERALPPTSWHPRALPNLIATRILVSMETGRPDARRRLAEAPVPAGG
FT      EEGVWWPALLLARARVAADDGDWEEALRLSRECGRWLFRRHWANPAMLSWRPLAAEACL
FT      KLGDVTEARRLRDEELFFADRWGTASARGIARLTTRRLFDDDDGDAVRRIREAAALLRD
FT      SPARLAYLWSRLSQAGAETAHGDTAAAARSWQAVARMTAAHPASRLATAARTLTVPSVP

```

```

FT      VATAPPTAVVPPGWRDLSEAEKDTVLLAARGHGNRQIAEQLAVSRRTVELRLSNAYRKL
FT      RIGGRKELYLLLEALEGPVADAS"
FT      CDS      117033..119816
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4V9"
FT      /note="transcriptional regulator"
FT      /transl_table=11
FT      /gene="nysRIII"
FT      /function="presumable transcriptional activator involved in
FT      regulation of nystatin biosynthesis"
FT      /product="NysRIII"
FT      /protein_id="AAF71780.1"
FT      /translation="MLLERENELARIRAALDAEAGDSSLLLLINGPLGSGRSALLRRIP
FT      ELAGDGTRVLRASAAWRERDFPFGIARQLFDHLLSGAGGAGPAERTAGAEHFSRLMDTG
FT      DRPTGTGPALEVSQAVLQGAQALLADASAERRLLILVDDLQWADGPSLRWLAHLTRRLH
FT      GLRALLVCTLADGDHGRYPLVREVAGAAHTVLRRLAPLSRDATRVLLAGPQGRPPQDAL
FT      VRVYEASRGNPLFLTAFRSALRATGRPPGGDHFGAVRELSPTVLRDRLAGHLRIQPQP
FT      VREVAVAVAALGDHSDPVLLAQLAGVDEIGFAGARRALVDAGLLARGRDVRFVHGVVRD
FT      AVDSLLTLDERERSHDDAADLLYRCGRPAEQVAGHLLAVVHPGRPWSEAVLRSAAHNAL
FT      RAGRPADAARYLRRALLHHRTQDGCRRARILVDLATAERALDPACVRHVSQAVALLDTS
FT      RDRAAAVLRIPPSLLAAPSPSAVELVRQAAAGLDEPGQRDEEGADELALRLEAWLRHSG
FT      HENPVELASSVLRMRMGARPPVDSVAERELVAVLLSAGALSGRLSAAEIADTGNRILE
FT      REPATAAHAHTPLPLVMLSFLVAESVQGVASWLASEQHTRRRYATGADDVLLTAERAFV
FT      LVTQGRPAAAREHVERALVMDAGDWSEPAVMMFAAVAFELRDPALSERILERIRDRRPA
FT      GLALTATGQMLQAAVDVHFGRRDALDTLLACGRRLETVGWRNSALLPWRPYAIGLHQR
FT      LGETDAALQLAEDELWAREWGATTNLGRALRLKGWLLQDEGLDLLRESVEILRASSYA
FT      TELARTLVVLGRRLPGGPAAEAVLREAAGIAAACGVPWLAERAELGLGSAIVPPVATLT
FT      PSERRVASLVSRLTNQAIATELGVSSRAVEKHLTSAYRKLGVSGRRELVNALPGR"
FT      CDS      120268..120900
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4V8"
FT      /note="putative transcriptional regulator"
FT      /transl_table=11
FT      /function="presumable transcriptional response regulator"
FT      /product="ORF4"
FT      /protein_id="AAF71781.1"
FT      /translation="MISAQTAPAGESVGPGLMASLDRDLTIKHANQEFRRRFFDDSAGDV
FT      CGRSFRDLMHPSVQQPLMRQFSRLIEGKRHRFASHVVAVGAQDAAFAGTLTASAVTGKT
FT      PDIAGILVLMDSGGAADAADAGVVTSSQKKFLTEIDARILEGIAAGLSTIPLASRLYLSR
FT      QGVEYHVTGLLRKLVRPNRAALVSRAYSMGILNVGTWPPKVVDFFIK"
FT      CDS      121589..122350
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4V7"
FT      /note="putative repressor"
FT      /transl_table=11
FT      /function="presumable transcriptional repressor (DeoR
FT      type)"
FT      /product="ORF3"
FT      /protein_id="AAF71782.1"
FT      /translation="MDAEGRRRDMLELIRRSAGADVVRRLAEFAVSKETVRRDLNVLEG
FT      HGLIRRRHGGAYPMVRPGSEAVFVSRTAQPIEESRIATAAAELLSEAEVTFIDEGFTP
FT      QLIADALPRDRPLTIVTASLPVVSFAFATSPQANVLLLGGRVRRGTATVDHWAVHMLSG
FT      FVIDLAFLGAEGISRRYGLTTPDPAVAEVKAQAIRVARRPVLAGVHTKFGTASFCRFGE
FT      VGDLETIVTGAGLPVAEAHRYHLMGPKVLRV"
FT      CDS      complement(122404..123468)
FT      /codon_start=1
FT      /db_xref="SPTREMBL:Q9L4V6"
FT      /note="putative transcriptional regulator"
FT      /transl_table=11
FT      /product="ORF2"

```

FT /protein\_id="AAF71783.1"  
 FT /translation="MAQDSGQTPRSLDHVDQALVHALQITPRASWTRIGSVLGLDAVTV  
 FT ARRWNRLVETGAAWISCHPAPVLAASGQGCLAFVEIDCAPGRLLDVARALAAAPHVVAL  
 FT SHVTGDRDLQLNVMARDPAMLSRWVTHDLAALDGVRAARTHLAGPVHTEGSRWRLRALG  
 FT RHQVARLAADASRHRTDTPAFVLDELQQLVTALSDVGRATYRALAEQCGAGPDTVRRR  
 FT VQRLFAADMLHARCEVARPLSEWPVTVSFWGQVPAARLREVTRRVTMREVRCLCASVIS  
 FT RHNHLVAVVRSLLDQAQRFEVRLAERAADLTVTERAVALWHMKHGGHLLDEEGYRVGVT  
 FT PLALWREPTDARRG"  
 XX  
 SQ

Sequence 123580 BP; 15426 A; 49056 C; 42187 G; 16911 T; 0 other;

gtatgaccta	tttcgccccg	tggcgtaagg	agtagccggc	aggtttcac	cgaaggtggg	60
gcgcgggagc	ggcgagttga	ccgtgaggtg	cgtgcggtgt	tcccggtcgc	gccggggggg	120
gggggcccgc	aggtggaccg	ccgtggtcag	cagggcgggg	gtgctgtgcc	agcggccggg	180
gaagaccgtg	acccggcgac	cgcccagcag	ggggcccggg	accaggagtc	gggcggtgaa	240
ggcgccctgt	gtcgggtcga	tgacgatctc	ggcctcggag	aagtccagtt	cgcgctgggt	300
caggggatac	cacgccttga	agacgctctc	cttggcgctg	aagagcagcc	ggtcccagtg	360
gacgtccgga	cgatgggccc	ccagggccac	gaggtgcggc	cgttccgagg	gcagggcgat	420
ggcggttcagg	acgccggccg	gcagcggacc	gttcgggttcg	gcgtcgatgc	tcaccgcggc	480
cgacagctcc	gcggggggaga	cggcgggcgg	gcggtagccg	gcgcagtgcg	tcagtgtgcc	540
gacgatgccc	ggcgggccact	gcggggcgcc	gcgcgcatgt	ggcagtatgg	cccgggtccg	600
gtggccgagc	cggcgccagg	cccggcgggc	gagatggcgg	acggtgggtga	actcgcgctg	660
ccgggactcc	acggcccggg	cgatgacctc	gcgttcagag	gagagcagcc	ggtcgccggg	720
gcgcggccgg	tcgtcgtacg	ccgcctcggt	ggcgaccgtg	gcgggcagga	tcagttcgat	780
caccgcatac	ctccggcgaa	cggtaagaac	aggggggttg	ttcggggcac	ggttccgtcc	840
ttgacggggc	gccgtggggc	gccgggtcag	ccggccggcg	cgcccgcggt	gggagtgtgg	900
gtgcgggctg	cgtgcagccg	ggcgtacagg	ccctgcgcgc	acaggagtgt	atcgtgggtg	960
ccctgctcga	cgatgcggcc	ggcgtccatc	acgacgatca	ggtccgcgct	gcggatcggt	1020
gacaggcggt	gcgcgatcac	gaagctgggt	cggcccgcgc	gcagggagtt	catggcgcg	1080
tggatcagga	cctcggtccg	ggtgtccacg	gagctgggtg	cctcgtccag	cacgaggacg	1140
gccggtctgg	cgaggaaggc	ccggggccac	gtcagcagct	gcttctcgc	ggcgtgacg	1200
gtgcccact	cgtcgtccag	cacggtgtcg	tagccctcgc	gcaggggtgc	gatgaagcgg	1260
tcggcacagg	tcgcgcgggc	cgccctcctc	atgtccgcac	ggcaggcacc	gggtgcgcgg	1320
tacgcgatgt	tctccgcgat	ggtgccggcc	aacagccagg	tgtcctggag	caccagccc	1380
aagcgggacc	gcaggtcgtc	gcgggtcatc	gtcgcggtgt	cggtgccgct	caggaggatg	1440
cggccggagt	ccggttcgta	gaagcgcac	aggaggttgc	cgaggggtgt	tttgccggcg	1500
ccggtggggc	cgacgatcgc	cacggtgctg	cccgggtcca	cggtcagcga	gaggttctcg	1560
atgaggggcg	tgtcggggga	gtagcggaag	gacacgtcgg	tgaactcgac	gcggccctcg	1620
gcgcgggccc	gcgtgccggg	ccggagcggg	tccggggcct	gctcgggggg	gtcagagcagg	1680
gtgaagacgc	gctgggcgga	ggcgatgccg	gactggagcc	ggcccggcc	cgaggcgatc	1740
tccacgatcg	gctgggtgaa	ctggcgggcg	tagaggatga	acgcctgcac	gtcaccgagg	1800
gtcagggtgc	cgtttatgac	cttccaggcg	ccgatgacgg	ccaccagcac	atagccgagg	1860
ttggcgacga	acatcatgac	cggttccatg	gcaccggagg	cgaactgcgc	cttgggcgca	1920
gcccggtaga	ccgcgtcggt	gcaggcgctg	aagcgtcctc	cgccgcgcgc	gcgcgggtcg	1980
aagcccttga	tcagcgcgat	accggtgcac	acctcctcca	catgggctgt	gaggggtccg	2040
ttcgcggacc	actgcgcggc	gtagtggggc	tgcgcgcgct	tgctgatccg	ggccgcgatc	2100
aacgccgaga	ccggcacgct	gagcagcatc	accacggcca	gcgacggcga	gatcaccagc	2160
atcagcacca	gcacgtcaa	cagcgagaag	atcgaggtga	tcagctcggc	gaggggtctg	2220
tggagggctc	gttggagggt	gtcgatgtcg	ttggtgggtg	ggctgagcag	ctcaccggcc	2280
ggctgccggg	cgaagtggcg	cagcggcagc	cgggtcagct	tctcccgggc	gtcgcggcgc	2340
agttcgtgga	tgggtgcgca	caccgcggac	gccaccagcc	ggccctgcgc	cagcatgaac	2400
agcgacgcca	cgacgtagag	cgccagcagg	accagcagca	gccggccgat	cgcggcgaa	2460
tcgatcccgg	gcgcggggcc	cgggacgccc	ccgagcacgc	cgtcggcgat	cagatcggtg	2520
acccggccga	gcagcagcgg	gccgaacgcg	ttgagcacga	tcccggccgac	gcccacgcac	2580
acggcgagtg	ccacggagcg	gcgggtgcgga	cgcagcaggc	cgacgagccg	acgtgccggc	2640
cggggcgcg	tccgctcctc	ctcaaggctc	tccggcgagg	ccatggggcg	cttcctcctc	2700
ggtcagctgc	gagagcgcg	tctcgcggta	ggtcgggctg	gtgcgcagca	gcacgtcgtg	2760
ggtgccctgg	gcgaccacgc	gcccgcggtc	cagcaccacg	atccgggtcc	cgtcgcggcc	2820
ggcgagatc	cgtcgcgcca	cgggtgatcac	cgtggcgccc	gcgggtgtac	gcaccagcgc	2880
ggtccgcagc	gccgcacg	tggcctggtc	gagcggcgag	aaacagtcgt	cgaagagata	2940
gatctccggc	cggcgacgca	acgcccgggc	gagggacagg	cgttggcgct	ggccgcggga	3000



gacattgccg	ccgccctggg	tgatctccgc	gtcagaggccg	tccggcatcc	gcgccacgaa	3060
gtcggccgcc	tgggcgaccc	gcagcgcctc	ccacagctcc	tcgtcgggtg	cgtccggccg	3120
cccgaagcgc	agattgctcg	ccacggtgcc	ggagaacagg	tacggccgct	gcggcacgaa	3180
cccgaagcgc	gcggcgagcg	tggccgcggt	cagctcgcgg	acgtcgggtg	cgccgacccg	3240
caccgcgccc	tcggtggcgt	cggccagccg	cagcaccaag	ttcaacaggg	tcgtcttgcc	3300
gctgccgggtg	ctgccgagca	cggcgatccg	ctcgcccggc	tcgacgggtc	ggtcgacgtc	3360
ccgcagcacg	ggctcctcgg	cgcccgggta	gcggtaccgc	gccgcgcaca	gttcgatccg	3420
gccggcgggc	ccgcgcaccg	gctgcggcgc	ggccggcggc	gccacgctcg	acccggtgtc	3480
caggacctcc	gcgatccggc	cggcacagac	ccgggcccgc	ggcaccgaca	ggaacacgaa	3540
ggcgagcatc	acgacggaca	tcaggatcag	cgagagatag	ctcaggaggg	cgctgagcga	3600
gccgatcggc	atccggcccc	cgctcgatccg	gtgggagcgc	gtccacagca	gggtcaggt	3660
gaaaccgttc	atcagagca	gcacgaccgg	cagcatcgtc	gcgatcagcc	gacccacccg	3720
ccgcgacacc	acgaggaacg	cgctcgttgg	ctgcgcgaac	cgcgcgcgct	cggtggtcgtc	3780
gcggacgaag	gaccggacca	ccgcacccc	ggtgatcgcc	tcgcgcagca	gccgccccag	3840
ccggtccagg	gtcagctgca	tccgcgcgta	cagggtgccc	atccggggca	gcagcaggcc	3900
gaagcagacc	gccaccacca	gcaccagcgc	caccagcagc	agtgccagcg	gaacgtcctg	3960
gcgcagcgcc	agcagcacgc	tgcccaggca	catcagcggc	gcgcagacga	cgatgccgaa	4020
gccggtctgg	gcgaggttct	gcacctgctg	cacgtcgttc	accgaccggg	tcagcaggga	4080
gggggtgccg	aaccggccga	tctcgcgggc	ggagaagtcc	aggatgcggc	ggaagagcgc	4140
ggaccgcaga	tcgcggccca	tcgccgtcgc	ggtccggggc	gccagcgcgg	ccgcaccgag	4200
cgcggcggcg	atctgcacca	gcgccaccac	gcccatacc	acacccagct	cggtgatgcg	4260
ccgcgcgtcg	ccgcgcacca	ccccctggtc	gataagtgcg	gcgccagcg	tcggcagcag	4320
caaagtgcc	aggatctgga	cgagttgaag	ggcgacgaga	gcggcggtgg	cccaggcgta	4380
ggggcgcgag	tgtgcccga	gaagtctcaa	cagcacggag	gaacacccc	gttgacggcg	4440
gcgtgcggcc	gcgcggtgga	cggagtcggc	cggcccgcgc	acccggtcat	tgcaccgcag	4500
tccgctccga	aatttcaacta	gtgttggggg	tggcaacgga	cttttgacgg	cccggtactc	4560
gtgaattccc	tagaaagccc	gggatgcgtt	gacagcatct	tccgcggctt	gcgagcgtgc	4620
gggtgtctat	cgtccggact	gcgtattcga	cggcaggctg	tggccgagct	caagcgtttg	4680
gacggtctct	ttcgtgttcg	agaagggttt	gccatgtcca	aacgagcgct	gatcaccgga	4740
atcacgggcc	aggacggctc	ctatctcgcg	gagcacctgc	tgtcccagg	ctaccagggtg	4800
tgggggtctga	tccgcggcca	ggccaatccc	cgcaagtccc	gggtcagccg	cctcgccctc	4860
gaactcgact	tcacgcagcg	ggacctgatg	gaccagggca	gcctggtctc	cgccgtcgac	4920
accgtgcagc	ccgacgaggt	ctacaacctc	ggcgccatct	cgttcgtgcc	gatgtcctgg	4980
cagcaggccg	agctggtcac	cgaggtcaac	ggcatgggcg	tgtgcgcgat	gctggaagcc	5040
atccgcgatg	tcagcggact	gtccacctcc	cgcacgggtc	gcccgcgcgg	ccagatccgc	5100
ttctaccagg	cgtccagctc	ggagatgttc	ggcaaggccg	ccgagacgcc	gcagcgcgag	5160
accacctctc	tccacccgcg	cagcccctac	ggcgcgga	aggcgtacgg	gcactacatc	5220
accgcgaact	accgcgagtc	cttcggcatg	tacgcgggtc	ccggcatgct	cttcaaccac	5280
gaatccccgc	gccgcggcca	ggaattcgtc	accgcgaaga	tcagcctggc	ggtcgccccg	5340
atcaagcaag	gcctccagga	caagctggca	ctcggaacc	tcgacgcggg	gcgcgactgg	5400
ggctatgccg	gcgactacgt	ccgcgccatg	cacctgatgc	tccagcagga	cgccggcgac	5460
gactacgtca	tcggcaccgg	gcagatgcac	tcggtgcgcg	acgcggttcg	gatcgcgctc	5520
gaacacgtcg	gcctgaactg	ggaggactac	gtcgtcatcg	accccgacct	ggtgcggccc	5580
gccgaggctg	aggtgctgtg	cgccgacagc	gccaaggccc	aggaccgcct	cggctggaag	5640
ccggacgtcg	acttccccac	cctcatgcgc	atgatggctg	attccgacct	ggcgcaggtt	5700
tcccgcaaaa	accaatacgg	cgacgtgctg	ctcgccgcca	actggtagca	gttctcaagc	5760
tttcgaaaac	tagtgaattc	ctgcccgaat	tccgacgaca	ctcgccacat	ggattcccc	5820
ggtgagtggc	gaatccagg	ggcgaatccg	aacgtaccgc	cgaaacggcg	tggagaagtc	5880
ggacgcccatt	cacgtgcggc	cgctcggtgt	cgagatgggt	tgagttgaga	tggacaacga	5940
acaaaaactc	cgggattacc	tcaagcttgc	gacggccgac	cttcgacgca	cccggcgggc	6000
cgtccacaag	ctggagtcgg	cggcccagga	accggtggcc	atcatcgcca	tgacctgtcg	6060
ctaccccggc	ggcgtccgca	gccccgaaga	cctctggcgc	atggctgagg	ccggcgagca	6120
cgcgctcacc	ccgttcccc	ccgaccgcgg	ttgggacctg	gaggcgctgg	ccgcgcgcgc	6180
gaccgcctcc	ggcggattcc	tgacgcagc	acccgacttc	gacgcggact	tcttcggcat	6240
ctcgccgcgc	gaggcggtcg	ccatggaccc	gcaacagcgc	gtcgtcctgg	aatccgcctg	6300
ggaggcgctc	gaacgcgcgc	gcacgcaccc	gacgtccgtg	aagggcagcc	gcaccggagt	6360
cttcacgcgc	gcgatggccc	aggactaccg	ggtcggcccc	gccgacggcg	ccgagggttt	6420
ccaactcacc	ggcaacaccg	gcagcgtgct	gtccggccgc	atctcctaca	ccttcggcac	6480
ggtcggcccc	gccgtcaccg	tcgacaccgc	ctgctcctcc	tccctcgctg	ccgtccacct	6540
cgccacccag	gcgctgcggg	ccggcgagtg	caccctcgcc	ctcgccggcg	gcgtcaccat	6600



catgtccggc	cccggcacct	tcatacgaat	gggcccgcag	ggcgggctct	ccgcccgcag	6660
ccgctgccgc	tccttcggcg	acaccgccga	cggcaccggc	tggggccgaag	gcgtccggcat	6720
cctcgctcgt	gaacggctgt	ccgacgccgt	ccgcaaccggc	cacgagatcc	tcgcccgtcgt	6780
ccgcggcacc	gccgtcaacc	aggacggcgc	ctccaaccggc	ctgaccgccc	ccaaccggccc	6840
ctcccagcag	caggatcatcc	agcaggccct	ggtcaaccgcc	cgactcgccg	ccggggacat	6900
cgacgtcgtc	gaggcgacag	gcaccggcac	caccctcggc	gaccccgtcg	aggcccaggc	6960
cctgctcgcc	acctacgggc	agaaccgccc	ggcggaccgg	ccgctgctgc	tgggctcggg	7020
caagtccaac	ctcagccaca	cccaggccgc	cgcggcgctc	gccggcggtga	tcaagatggg	7080
catggcgatg	cggcacggca	ccctgccgcg	caccctgcac	gccgaggagc	ccaccacca	7140
cgctcgactgg	tcgcaggggc	ccgtgcgggt	gctgaccgac	accaccgact	ggcccgcac	7200
cggggcgccg	cgcgcgcgcg	ccgtctcctc	cttcggcatc	agcggcacca	acgcccacac	7260
catcatcgag	caggcccccg	aaccgcagcc	cgaggacgcc	gcgaccgcgc	aggacgcgc	7320
cgcggcagc	acgcgggcca	ccgccccgt	agtgcgcggc	gtcgtaccgg	tcctgctctc	7380
cggccgcacc	ccggacgccc	tgcgcggcca	ggcgcgggcc	ctgcgcgcgc	ccctcgacac	7440
cggccggcgg	cccagcctgc	tcgacctcgc	acactccctc	gccaccaccc	gcgcggggtt	7500
cgagcaccgc	gccgtcctcc	tcgccaccga	ccaccgcgcc	ctgaccgacg	gcctcaccgc	7560
cctcgccgac	gccgacgacc	cggccgcgcg	ccccgcctgg	atcaccggca	ccaccggggc	7620
cgagaccggg	ctcgccgtcc	tgttcaccgg	ccaggggcgc	caacgcctcg	gcgcgggacg	7680
ggaactcgcc	gcccgtttcc	cggcggttcg	caccgccttc	gacgcggcgc	tcgacgcctt	7740
caccccgcac	ctcgaccgcc	ccctgcgcga	ggtcctgtgg	ggcaccgacg	ccgcctgct	7800
cgaccgcacc	gcatacgccc	agccggccct	cttcgcgcgc	gaagtggcgc	tctaccggct	7860
gatcgaatcg	ttcggcgctc	gcccgcacca	cctcgccggc	cactccgtcg	gcgagatcgt	7920
cgcgcgcac	ctcgccgggg	tcctctccct	ggcgcgcgcg	gccaccctcg	tcgcgcggcg	7980
cggtcgcctg	atgcaggcgc	tgcgcgacgg	cggggcgatg	atcgccgtcc	aggcgctcga	8040
agccgacgtc	gccccgcctg	tcgcgcggca	cgaggaccag	gtcgcgatcg	ccgcgcgcaa	8100
cggccctcc	ccgctcgtcc	tgtccggcgc	cgaagccacc	gtcaccgcgc	tcgccgaaca	8160
gctcgccgcc	gacggccgca	agaccgcgcg	gctgcgcgtc	tcgcacgcct	tcactcgcc	8220
gctcatggag	ccgatgctcg	acgccttccg	cgcgcgtcgc	gaagacctca	cgctccagcc	8280
gccgctcctg	ccggtcgtct	ccaacctgac	cggcaagccc	gccaccgtcg	cccaactcac	8340
ctccgcccgc	tactgggtcg	accacgtccg	gcacgcgcgc	cgttcgcgcg	acggcatcga	8400
ctggctcgcc	cggcacgcaca	ccaccgcctt	cctcgaactc	ggccccgacg	gcgtgctgtc	8460
cgccatggcc	caggactgcc	tggacgcgcg	cgacgcagac	gccgtcacc	tcgccgcctt	8520
gcgcgcgggg	cgcggcgagg	agcacaccct	caccaccgcc	ctcgccgggtc	tgcacgtcca	8580
cggcgccacc	ctggactgga	ccggctgctt	cgcgggcacc	ggcggcccgc	gcaccgacct	8640
gccgacctac	gccttccagc	gccgcgcgta	ctggcccaag	gccctccaga	gcggcaccgc	8700
cgacctgcgc	tcggtcggcc	tcggtgccgc	ccaccaccgc	ctgctctccg	ccgcctgctc	8760
cctcgccgac	gcaggcggca	ccctgctcac	cggccgcctc	tcgccgcaga	ccaccacctg	8820
gctcgccgac	cacaccgtcc	gcggcaccac	cctgctgccc	ggtaccgcct	tcctcgaact	8880
cgcctccgc	gccggcgacg	aggtcggctg	cgaccgcgtc	gaggaaactca	ccctcgccgc	8940
accgctcctg	ctgcccgaac	agggcggcgt	ccaggctccag	ttgtggatcg	gcaaccccca	9000
cgtgtccggt	cgcgcgacgg	tcaacgtcca	cgcggcggcc	gacaccggcg	acgacacccc	9060
ctggagccgc	cacgcccaccg	gcgtcctcac	caccgcgcac	gcctcccgcc	agctcccggc	9120
ttcgtccgag	cagggcggca	ccccctcgc	cggcgaccgc	caccccgccc	tcgacggggc	9180
ccagtggccc	ccggccggcg	ccgaaccgct	gccgctggac	ggccactacg	accgcctcgc	9240
cgacggcggc	ttcggtctacg	gcccggctctt	ccagggcctg	cgcgcgcgct	ggcgcggcgg	9300
cgacgtcgtc	tacgccgagg	tcgagctgcc	cgaggccggc	cggtcgcgacg	ccgaggcggtt	9360
cggcctccac	cccgcctcgc	tcgacgcgcg	cctgcacgcc	gcgccttca	ccggcctcgg	9420
cgaacgcggc	cggggcgggc	tgcggttctc	ctgggagggc	gtctccctcc	acgcggcgcg	9480
cgcaccacc	ctccgcgtcc	gcctgacccc	ggtcgcgcgac	gacgcgctcg	ccctgaccgt	9540
cgcgcgacgg	accggcgcg	ccgtgctgtc	cgtcgactcg	ctcgtcctgc	gcagcgtggc	9600
gacccaacag	ctcgacacgg	ccgccgcgct	cgcggcggtg	gccctcttcc	gcctcgactg	9660
gacccccgtc	cagccgaccg	ccaccgaccc	cggggccgctc	gccctcctcg	gcgcggaccc	9720
cttcggcctg	ctcaccacag	ccggattcgc	cgacgccccg	gcatacccg	acctcgccgc	9780
cctcgccgcg	gcggacggcc	cggtcggcgc	caccgtcgtg	ctgtccctcg	ccggcaccgg	9840
ggacgacgcg	gccgacccgg	cccgtccgcg	acaccgctgc	gccgcggagg	ccctcgccgc	9900
cgtacagacc	tggctcgacc	accatgagcg	cttcgcggcc	gcccgcctgg	tcctcgtgac	9960
ccgtggtgcg	acggtcgggc	gtgatgttgc	tgtgctgctg	gtgtgggggtc	tgggtgcgttc	10020
ggcgcagtcg	gagaatccgg	ggtgttttgc	tctggtcgat	ctggatccgg	atggtgcggg	10080
gggtgcggct	gcgctcgtcg	ctgcgttggg	cagtgtgtgag	ccgcagcttg	cgggtgcgcg	10140
tgatgtgttg	cgggtcgcgc	gtctggtgcg	gcggccgctc	accgaggtcg	gtgcgggtgc	10200

tgatggcacc	ggggatggcg	tcggggatgg	ctctgggtgtg	tcgttctcgg	gtgaggggtgc	10260
ggtcctgggtc	actgggtggtg	cgggtgggtct	gggtgcggtg	ttggcgcgctc	atctgggtggc	10320
cgagtatggg	gtgcggggatc	tgctgttggt	cagtcgcagt	ggtgaacgtg	ccgtgggtg	10380
tggggagttg	gtggcgggagc	ttgcggtgt	gggtgcgcg	gtgcggttg	ttgctgtgta	10440
tgtgaccgat	cgtgcccgcg	tgggtggagtt	gggtggcggg	catgcggtgt	ccgcggtggt	10500
tcatgcggtc	ggtgtgctgg	atgacggcat	gggtgggtg	ttgaccgggg	agcggttgtc	10560
cgcggtgctg	cggccgaagg	tggatgctgt	ctggcatcta	catgaggcga	cccgcggcct	10620
ggacctggac	gcgttcgtcg	tcttctcttc	cctcgccggg	gtcttcggca	gtcccggcca	10680
ggccaactac	gcggacgcga	acgccttctc	ggacgcgtg	atgacgcggc	gccgggcgga	10740
gggactgccc	ggcctgtcac	tcgcatgggg	accgtgggag	cagtcggggc	gaatgacggg	10800
cacctgacg	gcagtcgacg	ccgaacgggt	ggcccgctcc	ggtgtcccgc	cgctctccgt	10860
ggcgagggc	ctggccctct	tcgacgctgc	cgtggcggg	accgacgcca	cctgcgttcc	10920
ggtcgcgctg	gacctccccg	tcttcgcgc	acggggtgaa	gtgccgcgc	tgctgaggtc	10980
ggtgatccgc	gtccggggcg	gccgagccgc	cgtcgccggc	tccgccaccg	cgggcaacct	11040
cgcacagcgc	ctgcgcgcgc	tggacgagga	cggccgcgac	gagatggtcc	tggacctcgt	11100
ccgcggtcag	gtcgccctcg	tcttcggcca	cgcgaccggt	ggcgacgtcg	acgcggccg	11160
tgccttccgc	gacctcggtc	tcgactcgct	gaccgcgctc	gaactgcgca	accgcctcaa	11220
caccgtcacc	ggcctgcgcc	tgccgcgcc	cctgggtcttc	gactaccgca	ccgtccggca	11280
cctcgccacg	tacgtcctgg	acgagttggt	gggcacggat	gccgaggtgg	cgaccgtgca	11340
gccggcgcg	ggtgcggtgg	cggacgatcc	gatcgctcatc	gtgggcatgg	cctgccgcta	11400
ccccggtggc	gtcagctccc	ccgaggacct	gtggcgcggtg	ctcaccgaag	gcaccgacgc	11460
cgtctcgggc	ttcccgaacca	accgtggttg	ggacgtcgaa	tccctctatc	acccggaccc	11520
tgaccaccgc	ggtacctcct	acacgcgctc	gggtgggttc	ctgcatgagg	cgggggagtt	11580
cgatccgggg	ttcttcggga	tgagtcgcgc	ggaggcggtg	gcgaccgatt	cccagcagcg	11640
ggtgtgtgtg	gagtcgtcgt	gggaggcgat	cgagcgggcc	ggtattgata	cggtaggttt	11700
gcggggtagt	cggacgggtg	tggtcgcggg	ggtgatgtac	agcgattaca	gcgcgatggt	11760
ggcgagtcgc	gagttcgagg	gtttccaggg	cagtgggagt	tcgccgagtt	tggcgctggg	11820
tcgggtggcc	tacacgttgg	ggttggaaagg	cccggcggtg	acggtggata	cggcggtgttc	11880
gtcgctcgtt	gtggcgatgc	actgggcgat	gcaggcggtg	cgtagtggtg	agtgtgggtt	11940
ggcgttggcc	ggtggtgtga	cggtgatgtc	gacgcctgcg	gtgtttgtgg	actttgctcg	12000
gcagcggggt	ttgtcgccgg	atggccgggtg	caaggcggtt	gcggatgcgg	ccgatggtgt	12060
gggctgggtc	gagggcgctg	gcgtgttggt	cctggagcgg	cagtcggacg	cgggtgcgcaa	12120
tggtcacgag	atthttggctg	tgggtcgggg	ttcggcggtc	aaccaggatg	gtgctgccaa	12180
tggtttgacg	gcgccccaatg	gtccgtcgca	gcagcgggtg	atccggcagg	cgttggccag	12240
tgggtggtctg	acggccgggtg	acgtggatgt	ggtggaggcg	catggtacgg	gtacgacgct	12300
cggtgatccg	atcgaggcgc	aggcgttggt	ggcgacgtat	gggcgggata	gtgagcctga	12360
gcggccggtt	ttggtgggtt	cgggtgaagtc	gaatctgggg	catacgaggg	ctgctgcggg	12420
tgtggcgggc	gttatcaaga	tgggtgttggc	gatgcggcat	ggtgtggtgc	cgcggaagtt	12480
gcatgtggat	gcgccttctt	cgcattgtga	ctggtccgag	ggtgcggtgg	agctgctcag	12540
tgagcaggcg	gcctggccgg	agacgggtcg	ggtgcgggcg	gcgggtgtct	cctccttcgg	12600
catcagcggt	accaatgtgc	atgtcatcgt	cagcaggcg	ccgggcgcca	aggcgatcgc	12660
cgcggccggg	gcggcgcggc	gcaegccggg	tgccgtgccc	gtgctgctct	cggggcggtg	12720
ccggagtgcc	ctgcggggcg	aggccgcccc	cctgctcgga	cacctccagg	cccgaaccga	12780
cgcgaactc	gtcgatgtcg	cactgtcggt	ggcgaccacc	cggctccgct	tcgagcagcg	12840
ggccgcccgtc	gtggcgcgag	accgcgacca	gctgatcgcc	tcgctggggg	cgctggccgc	12900
cgaccgcccc	gaccccgccg	tcgtcgaggg	cgaggccgcc	ggacgcggcc	ggaccgcggt	12960
gctgttcacc	ggacagggca	gccagcgggc	cgccatgggg	cgtgaactcc	acgaggtgca	13020
gccggagttc	gccgcggcgt	tcgacgcggt	gtgtgcccgtt	ttcgaccgcg	tggtggaccg	13080
gccgctgcgc	gagggtggtg	tcgccgagga	cggcagcgac	gaggccgcac	tgctggacga	13140
gaccggttgg	acgcagccgg	ctctgttcgc	cgtcgagggtg	gcgctgttcc	ggctggtgga	13200
gagttggggg	gtccgtccgg	acttcgtggc	cggccattcc	atcggtgaga	tcgcggcggc	13260
gcactgcgcc	ggggtgctga	cgttggagga	cgcctgccgt	ctggtggccg	cgcgggcgac	13320
gctgatgcag	gcgctgccga	ccggcgggcg	gatgatcgcg	atccaggcca	ccgaggacga	13380
gatcgcgggc	cacctcgacg	acacgggtggc	gatcgccggc	gtcaacgggg	cgcagtcggt	13440
ggtgatctcc	ggtgacgagg	aggccgccga	aacgatcgcc	gccacgttcg	ccgaacgcgg	13500
gcgcaagacc	aagcggctgc	gggtgagcca	tgccttccac	tcgccgcgga	tggacgggat	13560
gctggacgcc	ttccggatcg	tcgccgaggg	gctgacctac	cgggcgcgcg	gcaccccgct	13620
cgtctccgac	ctcaccggcc	ggcgcgccga	cgatgcggag	gtgtgcaccg	cggagtactg	13680
ggtgcggcac	gtccgagagg	ccgtgcgggtt	cgcgcactgc	gtgcggacgc	tgcgcgacgc	13740
cggggccacc	accttccctg	aactgggctc	cgacggcgctg	ctgaccgcga	tggccgagga	13800

caccctcggt	gacgaccacg	acgccgaact	ggtgccgatg	ctgcgcgcgc	ggcgcgcgca	13860
ggaactggcc	gcggccaccg	ccctggcccc	cctccaggtg	cgcggcgtgg	acgtggactg	13920
ggcggcgtag	ctgcgcggca	ccggcgcccc	acgcaccgac	ctgccgacct	acgccttcca	13980
gcacgcgtac	tactggccgc	agctgccgac	cccggccgcc	gccctcgccg	ccgccgateg	14040
cgccgaccag	cagctgtggg	ccgctgtgga	gcgcggcgac	gcccgcgaac	tcgccgacat	14100
cctcggcctg	ggcgaacagg	acctcacgcc	gctggactcc	ctgctgcccc	ccctcacctc	14160
gtggcggcgc	ggcaaccagg	agaagcacct	cctggacacc	ctgcgctacc	gcgtggagtg	14220
gacacgactg	agcaagccga	ccgccccggt	cctcgacggc	acctggctgc	tggtcgccctc	14280
cgacgccacc	gcggccgacc	agccagccct	cctcgacggc	ctggccgacg	ccctcggttc	14340
gcacggcgcg	cgggtgcgtc	gcctgcttct	ggacgactcc	tgcgcggaac	gcgcggtgct	14400
cgccgaacga	ctggcgcgga	ccgccgacgt	ggacgcgcgc	acctagggtg	tgtccgtgct	14460
gcgcgtcgac	gagcgggacg	ccgacgactg	cccgcgcgtc	acctcgcgac	tggcgtgac	14520
cgctcgcttc	gtccaggccc	tcgccgacac	cggcgcgccg	ggccggctgt	ggaccgccac	14580
ccgcggcgcc	gtctccacca	accccgcga	cccggtcacc	caccccgctc	aggccgctgc	14640
ctggggcctg	ggccggggcg	tcgccttgga	gcacccacgg	ctgtggggcg	gcctggtcga	14700
cctgccgcag	gtcttcgacg	agcgggcccg	acagcggctc	gccgggatcc	tcgccgtcaa	14760
ggacgcaccg	gacggcgagg	accagggtgg	gctgcggggc	accggagtct	ccggccgcgc	14820
gctcgtccgc	cacaccgtcg	aagcgtgccc	cacggccgcg	gagttcaccc	ccaccggcac	14880
tgtcctgatc	actggtggca	ccggtggcct	gggcgcgcag	gtcggccggt	ggctggcccc	14940
cgccggcgcc	cagcacctcg	tcctgaccag	ccgcgcgggc	ccggacgcgc	cgggcgccgc	15000
cgaactccgg	gccgaactgg	agggctacgg	gccgtcggtg	tcgctcgctg	cctgcgacgt	15060
cgccgaccgg	gacgcgctcg	ccgcgcctct	caccgcactg	cccgaaggac	tgccgctgac	15120
cggtgtcgtg	cacaccgcag	gcgtcggcca	ctacggcccc	ctggacaccc	tgagaccgcg	15180
caggttcgcc	ggcctcacgc	ccgccaaagt	cgccggcgcc	gcccacctcg	acgccctgct	15240
cgccgaccgc	gaactggact	tcctcgctct	cttcggctcc	atcgccgggtg	tcggggcgag	15300
tggcaaccag	agcgctacg	gcgcgcgcaa	cgctacctc	gacgcgctcg	ccctgcaccg	15360
ccgcgcgccg	ggcctcgccg	cgacctccgt	cgctggggc	ccgtggggcc	aggccggcat	15420
ggccgcgcgac	gatgccgttt	ccgagaccct	gcgcgcgacg	ggcctcgggc	tgctcgaccc	15480
ggccccggcc	atgaccgagc	tgccgcgcgc	cgctcgccgg	caggacgtca	ccgtcacctg	15540
cgccgacgtg	gactggcagc	gctacgcacc	gctgttcacc	tcgcgccggc	ccagcgccct	15600
gatcgccggc	ctgcccaggg	tcgcgcgcct	cgccgcgcgac	gagcgcaccg	agcaggacgc	15660
caccggcgcc	tcgaggtcg	tcaccgcgct	ccgcgcctcg	gccgaacccg	agcaactgcg	15720
cctgctgacc	gacctcgtcc	gcaccgagtc	cgccaccgtc	ctcgccacac	gctccgcccga	15780
cgccgtgccc	gagggccgcg	ccttcgcgca	cgctcggttc	gactcgctga	ccgcggtcga	15840
gctccgcaag	cgctggggcg	ccgcgaccgg	gctgtccctg	cccagcacca	tggtcttcga	15900
ctaccgcaca	ccgctggaac	tcgcccagta	cctcgggggc	gagatcctcg	gcgcggtgct	15960
ggaagtgcgc	ggcccggctg	ccaccggcgg	cgccgacgac	gagccgatcg	ccatcatcgg	16020
catggcctgc	cgcttccccg	gcggcgctcag	ctccccgga	cagctgtggg	acctggtcgc	16080
ctccggcacc	gacgcgatca	gcgagttccc	cgtaaccgcg	ggctggcaga	ccgggcacct	16140
cttcgaccgc	gaccccgacc	ggcccggcac	cacctactcc	acctaggggc	gcttcctcca	16200
cgaggccgac	gagttcgacc	ccaccttctt	cggcatctcg	ccccgcgagg	cgctgggtcat	16260
ggacccgcag	cagtcgctcc	tgctggagac	cacctgggag	tccttcgagc	gcgcggggat	16320
ccgcccggaa	accttcgcat	ccacctgac	cggcaccttc	gtcggtcca	gctaccagga	16380
gtacggcctg	ggcgcgggcg	acggcaccca	gggccacatg	gtcacccggc	gcagccccag	16440
tgtgctctcc	ggccgactgt	cgtacgtctt	cggtctggaa	ggcccggcgg	tcaccgtcga	16500
caccgcctgc	tcgtcctcgc	tcgtggcgct	gcacctggcc	tgccagtcgc	tgcgcaaccg	16560
cgagagcaac	ctggccgtcg	ccggcgggcg	cacgatcatg	acgacgccc	acctgttcac	16620
cgcgcttcagc	cggcagcgcg	ccctcgccaa	ggacggccgc	tgcaaggcgt	tctccgacga	16680
cgcgacggc	atgacgtcg	ccgagggcgt	cggcgtcgct	ctcgtcgagc	ggctctccga	16740
cgcgacgcgc	aacggccacc	cggtcctggc	cgctcctcgc	ggctccgcca	tcaaccagga	16800
cggcgcctcc	aacggcctga	ccgcgcccaa	cgccccgtcc	cagcagaggg	tcacccgcca	16860
ggccctcgcc	aacgcccgcg	tcgcgcgccg	ggacatcgac	gccctggagg	cgcacggcac	16920
cggcacaccg	ctcggcgacc	ccatcgaggc	ccaggcactg	ttcgccacct	acggccgcga	16980
ccgcgacccc	gagagcgcg	tgctgctcgg	ctcggtgaag	tccaacatcg	gccacaccca	17040
gtccgcgcgc	ggcatcgcca	gcgtcatcaa	gatggtcatg	gcgtgcgcgc	actccgaact	17100
ggccgcgacc	ctgcacgccg	acgcgcgcgt	ctcgacgtg	gactggctcg	ccgggacggg	17160
ccggctgctg	acctaggcgc	gcgcctggcc	ggagaccggg	cgcccgcgcc	ggcccgcggt	17220
gtcctcgctt	ggcatcagcg	gtaccaacgc	ccatgtcctg	ctggagcagg	cgcccgctcg	17280
ggacacccc	gcccaggagc	ggcccgccgt	ggcgccgggt	cgatcgccg	ccggcgctcg	17340
cccgtgggtg	gtcaccgccc	gcagcgccgc	cgccctgcgc	ggccaggccg	agcgccctct	17400

cgcgacgccc	gaaaccgctcg	gaaccgcacct	gccggcccgcc	ggaccgctcg	acatcggcct	17460
gtcgctggtc	tccgcgcgcg	cccgtttcga	gcaccgtgcc	gtcgctcgcc	cgcccgcggg	17520
caccgacccg	ctggcgcgcg	tgcgcgcctg	cgcgacggac	gggcccctcg	ccgtggctcg	17580
ccgtggcgtg	gcggacgtcg	agggctcggac	ggtgttcgtg	ttccccggtc	agggttcgcg	17640
gtgggtgggg	atgggggtccc	aactccttga	tgagtccggc	gtgttcgcgg	agcggattgc	17700
cgagtgtgcg	gcggcactcg	ccgagttcac	cgactggctg	ctggctcgatg	tgctgcgggg	17760
tgtgggtggg	gcgccgctcg	tggagcgggt	cgatgtgggt	cagccggcgt	cgttcgcggg	17820
gatgggtgctg	ttggctgcgt	tgtggcgctc	ccgtgggtgtg	ttgccggatg	cggtgggtggg	17880
gcattcgcat	gggtgagatcg	ctgctgcggg	ggtgtcgggt	gcgctgtcgt	tgccgggacgg	17940
ggcgcggggtg	gtggcgctgc	ggagtcaggc	cattggctcg	gcgttggcgg	ggcggggagg	18000
gatgatgtcc	gtcgcgctgt	cggtggacgt	gctcgaaccg	cggttggctg	agttcgaggg	18060
gcgggtgtcg	gtggccgcgc	tcaacggccc	gcgctccgtc	gtggctgcgc	gcgagcccga	18120
ggcgctggac	gcgctgcacg	cccggctgac	cgcgcacgac	atccggggcc	gccggatcgc	18180
ggtggactac	gcctcgact	cgcaccaggt	cgaggacctg	cacgaggaac	tgctggaggt	18240
gctggcggag	ctggcgccgc	gcacgtcggg	ggtgccgttc	ttctcgaccg	tgaccggcga	18300
ctggctggac	accgcgcgga	tggacgccgg	ctactgggtc	cgcaacctgc	gcggacgggt	18360
gcggttcgcg	gacgcgggtg	cggacctgct	ggcgccggag	taccgcgcat	tcgtcgaggt	18420
cagctcgcat	ccggtgctgt	cgatggcggt	gcaggaggcg	atcgacgagg	ccggcgtgcc	18480
ggcgcgcgc	gccggcaccc	tgcgcgcgga	ccagggcggc	accgaccgct	tcctgctgtc	18540
ggcgcgcgag	gtcttcgtgc	gcggtgtgga	cgtggactgg	gcggggctgt	tcgaggggac	18600
cggtgcgtcc	cggatcgacc	tgcccaccta	cgccttccaa	cacgaacacc	tgtgggcccgt	18660
cccgcgccgc	ccggaggccg	tcgcccgcgc	cgaccgggac	gacgcggcct	tctggaccgc	18720
ggctgaggac	ggtgacgtct	ccgcgctgac	cgccgcgctc	ggcaccgacg	aggactccgt	18780
cgcgcgcgtg	ctgcccgcgc	tgacctcctg	gcgcgcgggc	cgccgcgacc	gctccaccgt	18840
ggacgcctgg	cgctaccgcg	tcgcctggaa	acccctcggc	ggcaccctgc	cgcaccctgc	18900
cctgaccggc	acctggctgc	tggtcaccgc	cgacggcatc	gacgacaccg	atgtggcagg	18960
ggcgttggag	acctacggcg	ccgaggtgcg	ccggctggtc	ctggacgagg	agtgcgtcga	19020
ccgcgcgcgc	ctgcccggagc	ggctggccgg	cgcggaggac	gtgaccggca	tcgtctccgt	19080
cctcgccgc	gccgagcggg	ccgacgcggg	accgggcacc	tccttgggtg	tcggcaccgc	19140
cctgaccgtg	gactgatcc	aggccctcgg	cgacgcggaa	atcgacgctc	ccgtatgggc	19200
ggtgaccgcg	ggcgcggtct	ccaccggccg	ggccgacgag	ctgaccgcgc	ccgtccaggc	19260
acaggtcacc	ggcatcggct	ggaccgcggc	gctggagcac	ccgcagcgct	ggggcggcac	19320
cctcgacctg	cccgcgcgcc	tcgacgcgcc	ggcgccccag	cggtctgcgc	ccgtgctgtc	19380
cggcgcacctc	ggcagcgacg	accagctggc	catccggccc	tcgggggtct	tcaccgcgcg	19440
catcgctgcg	gccgaggcca	ccgcgcggcg	gcccgcgggc	acctggacgc	cgcgcggcac	19500
cacactggtc	accggcggtc	ccggcaccct	cgccccgcac	ctcgcccgtc	ggctggccca	19560
acgcggcgcc	gagcacctgg	tcctgatcag	ccggcgcggc	acggccgcgc	cgggcgccgc	19620
cgaactcgtc	gcggaactgg	ccgagtcggg	caccgaggcg	accgtcgccg	cctgcgacat	19680
caccgaccgc	gacgcggctg	ccgcgctgct	ggccgacctc	aaggccgacg	ggcgcacctg	19740
ccgcaccgtc	gtgcacaccg	ccgccaccat	cgagctgcac	accttggacg	ccaccacctc	19800
ggcggaactc	gaccgggtgc	tgacgcgcaa	ggtcaccggc	gcccagggtc	tcgcggaact	19860
gctcgactgac	gaagagctgg	acgacttcgt	cctgtactcc	tcaccgcgcg	gcattgtggg	19920
cagcggcgc	cacgcgcctc	acgtcgccgg	caacgcctac	ctcgccgcgc	tcgccgagca	19980
ccgcgcgggc	aacggactgc	ccgcctctgc	gctgtcctgg	ggcatctggg	ccgacgacct	20040
caaactgggc	cgggtcgatc	cccagatgat	ccggcgagc	ggcctggagt	tcattggaccc	20100
gcagctggcc	ctgagcgggc	tgacgcgggc	gctggacgac	aacgagaacg	tgctcgcggt	20160
cgccgacgtg	gactgggaga	cctaccaccc	cgtctacacc	tcgggcccgc	ccaccccgct	20220
cttcgacgag	gtgccggagg	tcgcgcggct	caccgcggcc	gccgagcaga	gcgcggggac	20280
cgctcgccgag	ggcgagttcg	ccgcgcgcgt	gcgcgcctcg	tcgacgcgcg	agcagcagcg	20340
caccctgctg	gagaccgtcc	gcaccgaggc	ggcgctccgtc	ctcgggctgt	cctccgcgga	20400
ggacctcacc	gaccagcggg	ccttcgcgga	cgctggcttc	gactcgctga	ccgcgcgtcg	20460
cctgcgcaac	cggctcgctc	ccgtcaccgg	cctgacgctg	ccctcgacga	tggtcttcga	20520
ctaccccaac	ccggcgcggc	tcgcgcgcta	tctgcacggc	gagctggccg	gcgcgccgtc	20580
cgcgcgcgc	ggcgccgcgc	ccgtcccgcg	cggcgccccc	gacgcgcgac	acccgatcgc	20640
gatcgctggc	atgagctgcc	gctaccccgg	cggggtcggc	tcgcgcgagg	acctgtggcg	20700
gatcgccctg	gacgaggtcg	acgcgatctc	cggcttcccc	gccgaccgcg	gctgggacgc	20760
cgagggcctc	tacgaccggg	accccgaccg	gcccggccgc	acctactccg	tcaggggcgg	20820
attcctgcgc	gacgtcgccg	agttcgaccc	gggcttcttc	gggatctcgc	cgcgcgaggg	20880
gctgtcgatg	gaccgcgacg	agcggtcctc	gctggagacc	gcctgggagg	cgttcgagca	20940
cgcgcggcatc	gacccggtcg	gccagcgcgg	cagccgcacc	ggcaccttcg	tcggcgccag	21000

ctaccaggac	tacgcctccg	gcgtgcccaa	cagcgagggc	tccgaaggcc	acatgateac	21060
cggcacgctc	tccagtgtgc	tgtccggccg	ggtgtectac	ctcttcgggt	tccagggccc	21120
cgcgctcacg	ctcgacaccg	cctgctcttc	ctccctggtc	gccatgcacc	tggcctgcca	21180
gtccctgcgc	aacggggaga	gtcgtctggc	cctggccggc	ggcgtcagca	tcatgtccac	21240
cccgatgtcg	tctgtcggct	tcagccggca	gcgcgccttc	gccgaggacg	gccgctgcaa	21300
ggcgtacgcg	gacggcgccg	acgggatgac	cctcgccgag	ggcgtcggcc	tgggtgtgct	21360
ggagcggctg	tccgacgccc	gcgccaacgg	gcaccagggt	ctcgccgtga	tccgggcttc	21420
cgcgctcaac	caggacggcg	cctccaacgg	cctgaccgca	cccaacggcc	cgtcccagca	21480
gcgcgtcatc	cgccaggcgc	tggccaactc	cgcggtggcg	cccgccgaca	tcgacgtcct	21540
ggagggccac	ggcacccgta	ccgccctcgg	cgaccccatc	gaggcgcagg	ccctgctcgc	21600
cacctacggc	caggaccgcg	ccccgaacg	gccgtgtgtg	ctcggtcctg	tgaagtccaa	21660
catcgccac	accagatgg	catccggcgt	cgcacgcgtc	atcaagctcg	tccgcgcctt	21720
ccaggaaggc	gtggtgcca	agtccctgca	catcgaccgg	ccctccaccc	acgtcgactg	21780
gtcctcgggc	gccatcgggc	tgtcaccga	acgcaccccg	tggcccagaa	ccggccggcc	21840
gcgcgcgcgc	gccgtctcct	ccttcggcat	cagcggcacc	aacgtccaca	ccatcctcga	21900
acaggccccc	gcggacgagg	cgcccacgcc	cgcgcacccg	ccgcgggacg	gcctggtgcc	21960
ggtcctgtct	tccggccgcg	gcgaggccgc	gctgcgcgcc	caggccgcgc	gcctgctcgc	22020
cttcgtcgag	gagcggcccg	aggccacct	caccgacctc	gcccactccc	tcgccacctc	22080
gcgcgcgcgc	ctggaacgcc	gcgcgcgcgt	catcgccgcc	gaccgcgaca	ccctgacccg	22140
cggcctgcgc	gccctgtccg	acggccggcc	cgaccccgcc	ctggtccagg	gcaccgcggg	22200
acgcggcccg	accgccttc	tgttcaccgg	acagggcagc	cagcgcgccg	gcattggccg	22260
cgaactccac	gaccgtacc	cgggtgttcg	cgacgcgcgt	gacgaggtgc	tggcccggct	22320
cgacgacgga	cgggacgggc	cgtgcgcgca	ggtgctgttc	gcgcgcgccg	actccgcga	22380
ggccgcgcct	ctggaccgga	ccggctacgc	ccagcccgcg	ctgttcgcgc	tcgaggtcgc	22440
gctgttccgc	ctgctgacgt	cctggggcct	gaccccgggc	tacctggccg	gccactccgt	22500
cggcgaactc	gccgcgcgcg	acgtcgccgg	cgtgctgtcg	ctggacgacg	cctgcaactc	22560
ggtcgcgcgc	cgcggccggc	tcacgcaggc	gctgcccag	ggcggcgaga	tggtcgcctt	22620
ggaggccgcg	gaggacgagg	tctgcccgt	cctggagggg	ctcaccgacc	gggtgtccgt	22680
cgcgcgcgtc	aacggggccg	ggtccgtggt	cgtgcgcggc	gtcgaggagg	acgtgtcctt	22740
cctcgccgac	ctcttcgcgc	ccgacggggc	ccggaccaag	cggctgcggg	tgagccatgc	22800
cttcactctg	ccgctgatgg	acgccatgct	cgacgacttc	gccgcgcgtc	cccgccgggt	22860
gacctaccac	ccgcccagca	tcccgttcgt	gtcgaacgtc	agcggcggcc	tggccaccgc	22920
cgaacagggt	cgcacgcccg	actactgggt	cgggcacgtc	cgcgcgcgcg	tccgcttcgc	22980
cgacggcatc	gactggctcg	ccaccagggg	cgacgtccac	accttcctgg	agctcggccc	23040
ggacggcgtg	ctcagcgcca	tggcccggga	gagcctcacc	gacccgtccc	gcacggcact	23100
gctgccgacc	ctgcgcggcg	accggcccga	ggaacctgcc	ctggtaaccg	cggtcgcgcg	23160
ggcccacgcg	cacggcgccc	gcgtcgactg	gagcgggtac	ttcgccgacc	acggcgcgcg	23220
ccggaccacg	ctgccgacct	acgcgttcca	acgcgagcgg	tactggcccc	acaccaccgc	23280
cgccacgagc	gccccacgcg	ccggatccgc	cctcgacgcc	gagttctggg	ccgcgcgcga	23340
gcgggacgac	gtgcgcgcgc	tcgcccgcct	cctggacgtg	gacgacgcca	ccgtcaccgc	23400
gatgggtccc	gcgctcaccg	cctggcgccg	cgccgcgcgc	gagcagaccg	aactggactc	23460
ctggcgctac	cgcgtcacct	ggaagccgcg	cggcgccgcc	accgcacccg	ccgccctcac	23520
cggccgctgg	ctcgtgctcg	tcccgcacga	ccaccaggac	cgtaaggacg	acgcgaccgc	23580
ggcctgggca	gccgacgtcg	agaccgccct	gggcaccacc	accgtccggc	tgacggtcac	23640
caccaccgac	cgcgcgcgcg	tggccgcccg	gatcaccgaa	gccgcgcggc	accagggcc	23700
gttcagcggg	gtgctgtccc	tgtgcgcgt	cgcacccggc	gacgcgggcc	accccggtgc	23760
gcccgcgcgc	ctcaccctca	ccaccaccgc	cgtccaggcc	ctcgccgacg	ccggcatcga	23820
cgcgcgcgtg	tggaaacgtc	ccgcgcgggc	cgtggccgtc	ggccgcgcgc	aacaggtcac	23880
cgcgcgcgaa	caggccgcgc	tctggggcct	gggcgcgcgc	gtcgccctgg	aactgccggc	23940
ccggttcggc	ggcaccctcg	acctgcccgc	caccctggac	ggccaggccg	cccgccgggt	24000
gcgcgcgggt	ctcgcggtca	ccgacggcga	ggacgcgggt	gccctgcggc	cctccggcgt	24060
cttctctcgc	cgcctggccc	acgccccggc	cggccccgac	accgccccga	ccgccttcga	24120
cccggcgggc	ggcaccgtcc	tgatcacccg	cggcaccggc	ggcatcggcg	gccacgtcgc	24180
ccgcgcctcg	gcccgcgacg	gcgccaccca	cctgctgttc	accagccgcg	gcggcccggc	24240
cgcgcgcggc	gccgacgcgc	tccgcgcgga	actggaggaa	ctgggcgcgc	gggtcacctt	24300
cgcgcgcctg	gacgcgcgcg	accgcgacgc	gctggccgcg	ctcctcgccg	aactgcccga	24360
cgacgccccg	ctgtgcgcgg	tgttcacacg	cgcggcgctc	gtcgaggacc	acgtcgtgga	24420
cgcgtcacca	cgggaagact	tcgcccgggt	gctgcgcgcc	aagaccgtcg	ccgcccacca	24480
cctgcacgag	ctgacgcgcg	acctggacct	cgcgcgttct	gtgctgttct	cctccacggc	24540
cggcgctcct	ggcgccgcgc	gacagggcaa	ctacgcgcgc	gccaacgccc	acctggacgc	24600

cctcgccgaa	caccgcccgt	cccacggcct	gaccgcccgt	tccgtcgcc	ggggcccgtg	24660
ggccgggtcc	ggcatgggtc	ccgacggcgc	cgaactcacc	gaccgggtac	ggcgccggcg	24720
cttcgaaccg	ctcgccccc	aaccggccgt	gcgcgccttg	ctgcgcgcca	tcgagaacga	24780
cgacaccacc	gtcgcgctcg	ccgacatcga	ctgggagcgc	ttccagcgcg	ccttcgcccgc	24840
ggtccgcccc	ctgcggttcg	tgcgcgacct	ccccgagacc	ggccggggcca	cccccgcgac	24900
cgccaccggc	gccgccaccg	gacctgcgca	gcaactcgcc	gaactgcccg	agcacgagcg	24960
cccggcagcg	gtcctggacc	tgtcgcttac	ccaggtcgcc	gccgtcctcg	gccacgccga	25020
cccgccgacc	gtcgaggacg	accacgcctt	ccgcgacctg	ggcttcgact	cgctgaccat	25080
cctggaactg	cgcaacgccc	tcaacgcccgc	caccggcctg	agcctgcccg	ccaccctggt	25140
ctacgacctg	cccacccccgc	gcgagatggc	ggacttcctg	ctcgccgaac	tcctcggcac	25200
cctgcccacc	gacaccgccc	cgaccgtcgc	cagcacggcc	ttccccaagc	tctcagcttc	25260
gttcgagcag	ggcggtaccc	ccttcgacga	cccgatcgcc	gtcatcgcca	tcggctgccc	25320
cttccccggc	ggcgtcacca	ccccggagga	gctctggcag	ctcctcgacg	agggccgcga	25380
cggcatacag	cgcttccccg	acgaccgcgg	ctgggacctc	gccgcgctgg	gcgccggcgc	25440
ctccgacacc	ctggagggcg	gcttcctgac	cggcgctcgcc	gacttcgacg	cccggttctt	25500
cggcatctcg	ccccgcgagg	cgctggccat	ggacccccag	cagcggtgc	tgctggagac	25560
cacctgggag	gcgctggagc	gggcccggcat	cgacccgacc	acgtcgcgcg	gctccaccac	25620
cggcgctctt	gtcggcacca	acggccagga	ctacccgacg	ctggtgcgcc	gctccgcctc	25680
ggacgtggcc	ggctacgtcg	ccaccggcaa	caccggccagc	gtgatgtccg	gccgcctgtc	25740
ctacgcgctc	ggcctcgaag	gcccggccgt	caccatcgac	accgcctgct	cctcctcgct	25800
ctgcgcctcg	cactgggccc	gcccggcgct	gcgcgcggcg	gagtgcgacc	tcgtggtggc	25860
cggcgcgctc	tcggtcatgg	ccagcccggga	ctccttcgct	gagttctcca	cccagggcgg	25920
cctggcaccc	gacgggcgct	gcaaggcggt	ctccgacgcc	gccgacggca	ccgcctggtc	25980
cgaaggcgct	ggcatcctcg	tcctggaacg	cctctccgcc	gcccgcgcga	acggccacca	26040
ggtcctcggc	ctgatccgcg	gcaccgcccgt	caaccaggac	ggcgcgctcca	acggcctgac	26100
cgcgcccac	ggcctctccc	agcagcgcg	catcgcccag	gactcgcccg	acggccgcct	26160
gcgccccgcc	gacatcgacg	cgatcgaggc	gcacggcacc	ggcaccaccc	tcggcgaccc	26220
gatcgaggcc	cgcgccttga	tcaccgccta	cggccgggac	cgggacgccg	aacggccgct	26280
gctgctgggc	accgtcaagt	ccaacatcgg	gcacaccacg	gccgcgcgcc	gtgccgcggg	26340
cgatcatcaag	atgctgatgg	cgatgcgcca	cggcaccctg	cccaggacgc	tgacagtggg	26400
caccccgctc	agccacgtcg	actggagcgg	cggcaccgtc	gcgctcctcg	acgacgcgcg	26460
gccctggcca	cggaccgggc	agccgcggcg	cgcggcgctc	tcgccttcg	gcgtcagcgg	26520
caccaacgcc	cacgtcgctg	tcgagcaggc	cccggaaacc	gaagcccccg	ccgccccggc	26580
cgcgagcccg	gcgcgggagg	ccacgcccac	cgctgtcccc	tgggtcgctt	ccggacgcag	26640
ccgggaagcg	ctccaggcgc	agctggaccg	gctcaccgcg	cacaccgccg	cccacccccgc	26700
gcgctcggcg	gcggacgtcg	gcccgtcgct	ggccaccgac	cgcacgctct	tcggcgaccg	26760
cgcgctgctg	ctcgccggcc	cggacggggg	gcgcgaggcc	gcccgcgccg	ccgcgccccg	26820
caccccgggc	cgcaccgcgt	tcctgttctc	cggacagggc	gcccagcacg	ccctcgaggg	26880
ccacgacctg	taccagcgct	tcgggtctta	cgcgcagcca	ctggacaccg	tcctcgccca	26940
gttcgacacc	gtgctggacg	tcggctgctg	cgcgcgctg	ttcgccgcgc	cgggcacccc	27000
cgaggccgcg	ctcctggacc	agaccggcct	caccagccc	gcgctgttcg	ccgtcgaggt	27060
cgcactgttc	cggctcgccg	agtcctggcg	gctgacgcgc	gacttcgtcg	ccggccactc	27120
catcggcgag	atcgccgcgc	cgcacgtcgc	cggggtgttc	tccttgaggg	acgcctgcac	27180
gctggtcgcc	gcccgcgcct	ccctcatgca	gcaactgccg	cgggacgggtg	cgatggtggc	27240
cctggaagcc	accgaggacg	aggtcgcgcc	gctgctcacc	gacggcgctc	cactcgccgc	27300
ggtcaacggc	ccccgctcgg	tggtcgctgc	gggcgcgcag	gacgcgctcc	gcgcggctgc	27360
cgaccggctc	gccgcgcgac	gccgcgcgac	ccgcggcgctg	acggtcagcc	acgccttcca	27420
ctcgccgctg	atggacccga	tgtcaccgga	cttcgcccgg	gtcgcgaggg	gcctgacctg	27480
ccacgagccg	cgcatacccc	tcgtctccac	cctcctcggc	gccccggccg	gcgccgaact	27540
gcgcaccccc	gactactggg	tgccggcacgt	ccgcgagacc	gtgcgggttcg	ccgacggcgt	27600
gcgcgccttg	cacgacgcgc	gcgcgggcac	cttcgtggag	atcgggcccg	acggcggtgt	27660
caccgcccgt	accagcgaga	ccctcgacac	cgtcgaggcc	ggcgcgcccc	ccgtcgctgt	27720
gcccgtccag	cgcgcgcgac	gcgcggcgga	cctcgccctc	ctggaggggc	tgccaccct	27780
gcacacccac	ggcaccggcc	cgctcctggc	cgcctacttc	gaggccaccg	gcggccaccg	27840
gaccgatctg	cccacctacg	ccttcacgcg	ggagcggtac	tgccccgaac	tcggcgaccc	27900
cgtcgccacc	gccccgcagg	acccggcgcg	ctggcgctac	cacgagacct	gggccccgct	27960
gccgggaccc	gagggggccc	cgcgggcccg	cgcgcctctg	gtcctcgctc	cggccgggaa	28020
ccgcgacacc	cggtggatga	cggccgctgc	cgcgcgctc	ggcgccgaca	ccgtcacggc	28080
cgaacccgac	gccctggccg	agcagttgac	ggcgccggcg	gacacaccct	ggcgcgctgt	28140
ggtgtcgctg	ctcgccgcgc	cgctccgagg	gctgcccgcg	gacggcgcc	ggcccgccgc	28200



cctcctcgcc	accctggacg	aggccggcgt	gcacgcgccc	ctgtggtgcg	tcacccgcgg	28260
cgccgtcgcg	gtcgcggggg	aggccccgac	cgccgtcggc	caggccgccc	tgtggggcct	28320
gggcccgggtc	gcccgcgtgg	accacccgga	cgcttcgggc	ggcctggccg	acctgcccgc	28380
cgacaccgac	gcgacgcgcg	ccgggtctgt	cgccgcgcac	ctggccgcgc	cgggcaccga	28440
ggccgagatc	gcggtccgcg	ccaccggcgt	ccacgcccgt	cgcttggtcc	gtacgcgggc	28500
cgcgcgcgac	ggtgccacct	ggctgccgac	cggcaccgtc	ctggttgctg	gcggcaccgg	28560
cggcaccggc	accatgggcg	gccggggcgc	ccgctggctg	gtccgcgagg	gcgcccgcga	28620
cctcgtcctg	accgccccg	acggcaccac	gaccgcgcgc	gacaccgagg	ccctgacggc	28680
cgaactggcc	gcgctcggcg	cccggatcac	cgctcgtggac	cacgacccca	ccgccccgga	28740
cggcttcgcc	gcgctcctcg	acggactgcc	cgacgacacc	ccgctcaccg	cggctcgtgta	28800
cgcgcgggag	gcccagcgcg	ccccgggcac	cgcggccgag	ctgtccgcgc	cactcgcgcc	28860
cgtcaccgcc	ctaggcgccg	ccctcaccgg	ccggccgctg	gacgccttcg	tcctcttcgg	28920
ctccatcgcc	gggtctctgg	gcgtgcgcgg	ccgggcgcgc	gaggccgcgt	ccggcgccct	28980
cctcgacgcc	ttcgcccgcg	cctgccgcga	ccgcggcacc	ccggcactgg	ccgtcgccct	29040
gggcgccttg	gcccagcctg	tcggcccgcg	cctcgccgcg	cacctgcgga	tgaacggcct	29100
gccggtgatg	gacgcggaca	ccgcactgac	cgccctcagc	cgggcccgtc	ccgacgggtc	29160
cgccgcgcgag	gcggtcgccg	acgtccgctg	ggagaccttc	gcgcccctcc	accacgaggc	29220
ccgcccgcacc	gccctgttcg	acgcccctgcc	cgaggcccgc	ggcgcgctcg	cggaggccgc	29280
ccgggaccgc	gcccagccgga	agaccgcgcg	cggcgactac	ggccgggtggc	tcgcccagca	29340
gcccgcgcgcg	gaccacgacg	ccatcctgct	ggcactgggtc	accgagaagg	ccgcgaccgt	29400
cctcgccac	gcccagccacg	acctgctcga	accgacctg	cccttcgcgc	acctgggctt	29460
cgactcgctg	accgcggctg	acctgcgcaa	ccagctcacc	gcggaaccg	gcctcaccct	29520
gcccgcacc	ctcgtcttcg	accaccccaa	cccggccgcg	ctcgccgcgc	acctgcgcgc	29580
ccaactcctc	ggcgaggcga	gcgactccgc	cgcaccgggtg	gcccgcgcgc	tcgcccctcg	29640
tggcgacgac	gacgcgatcg	tcactcgtcg	catggcctgc	cgtacccccg	gcgggggtcac	29700
ctcgcccag	gacctgtggc	agctggctcg	cgacgaggtc	gacgcggctc	gcgacttccc	29760
gaccgaccgc	ggctgggacc	tggccgcgct	cgccggcgac	ggaccggggc	gcagtgccac	29820
cgcccagggc	ggattcctct	acgacgccac	cgacttcgac	cccggcctgt	tcggcatctc	29880
gcccgcgcgag	gcccctggtga	tggacccgca	gcagcggatc	ctgctcgaaa	cgtcctggga	29940
ggccctggag	cgggcccggca	tcgacccggc	gacgctgcgc	ggcagcggca	ccaccggcgt	30000
cttcgtcggc	ggcggtccg	gcgactaccg	gcccgcggag	gaggccgggc	agtggcagac	30060
cgcccagtc	gccagcctgc	tctccggtcg	cctcgccctac	accttcggca	tcaggggccc	30120
caccgtgtcg	gtcgacaccg	cctgctcctc	gtcgtgggtc	gcgctgcacc	tggccgcgca	30180
ggccctgcgc	gcccgcgaat	gctcgatcgc	gctggccggc	ggcgtcaccg	tgatggccac	30240
cccgtggggc	ttcgtcgaagt	tcagcgccca	gggcgcctcg	tcgcccggacg	gccgctgccg	30300
cgccttctcc	gacgacgcca	acggcaccgg	ctggtccgaa	ggcgtgggca	tgctcgtcgt	30360
cgaacgggtc	tcgacgccc	gcccgaacgg	ccaccgcgtc	ctcgccgtgc	tcgcgcgctc	30420
cgccatcaac	caggacggcg	cgtccaacgg	cctgaccgcc	cccagcggcc	ccgcccagca	30480
gcggtcatc	cgccaggccc	tcgccaacgc	ccgactgcgc	cccgcgcgaca	tcgacgccgt	30540
cgaggcccac	ggcaccggca	ccaggctcgg	cgaccccatc	gaggcccagg	ccctgctcgc	30600
cacctaaggc	caggaccgcg	agcggcccgt	gctgctcggc	tcgctcaagt	ccaacatcgg	30660
ccacaccag	gcccgcctcc	gcgtcggcgc	cgtcatcaag	atggtcctcg	ccatgcagca	30720
cggcgaactg	ccgcgtctcc	tgtacgcgga	gaacccctcg	tcgcacgtgg	actggaccgc	30780
cggccgcgcg	cacctgctca	ccgccaggac	cccgtggccc	gactccgggtc	ggccgcgcgcg	30840
cgccgcgcgtc	tcctccttcg	gcgccagcgg	caccaacgcc	cacgccatcc	tggagcagcc	30900
gcccgcgcgag	gaactccccg	cgcgccccgc	ggacgacggc	gccccgctgc	cgttcttgc	30960
ctccggccgc	tcgcagaacg	ccctgcgcgc	ccaggcccgc	cgactcctgg	cccgcctcac	31020
cgcccacccc	gacaccgggg	ccgcgcacct	ggcgtactcc	ctggcgacca	cccgggcccgc	31080
cttcgagcac	cgggcccgcga	tcaccgccac	cgaccacgac	ggcctccgca	ccggcctgac	31140
cgccgtcgcc	gagggcacca	ccgccccgca	caccgcgcaa	caccacctcc	agggcaccgg	31200
aaagcgcgc	gtgctcttct	ccggccaggg	ctcccagcgc	ctgggcatgg	gccgcgaact	31260
gcacgagcgc	caccgggtgt	tcgcccaggc	gttcgactcc	gtactggccc	gcctcgacga	31320
ccggctcgac	acccccctgc	gggacgtcgt	ctggggcacc	gacgaggagg	cgctgcacgc	31380
caccgggaac	acccagcccc	ccctgttcgc	cgtcgaagtc	gcgctctacc	gcctgatcga	31440
atcctggggc	gtgcggcccc	acttcgtggc	cggccactcc	gtcggcgagc	tcgcccgggc	31500
ccacgtcgcc	ggggtgctct	ccctggacga	cgccctgcgc	ctggtcgccg	cccgcgcgcg	31560
cctcatgcag	cgctcccg	ccggcggcgc	catgatcgcc	gtcgaggcca	ccgaggacga	31620
ggtcaccgcc	ctcctcaccg	acggcgtgtc	cctcgccgcg	gtcaacggac	cgaccgcctg	31680
ggtcctctcc	ggcgcgggcg	acgccgtgac	cgccctgggc	caggcgctgg	ccgaacgggg	31740
ccaccgcacc	acccggctgc	gggtcagcca	cgccttccac	tcgcacctca	tggaccgat	31800

gctggcggac	ttccgcaccg	tgcgcgaggg	cctggaatac	cacccgcgcg	gcatccccgt	31860
ggtctccaac	ctcaccgggg	acgtcgccga	cgcgcccgac	ctgtgctccg	ccgactactg	31920
ggtgcgccac	gtccgcggca	ccgtacgggt	cgccgacggc	gtgcgacca	tggccgaccg	31980
cggcgtgcac	ctcttccctg	aactcggccc	ggacgcctg	ctgtcggcca	tggcccgcga	32040
gtgcgacccg	gacgcgcgtg	tgtcccggc	cctgcgcgcg	aaccgcgacg	aggacgagac	32100
gctggtcggc	gccgtcgcgc	gactgcacgt	ccacggcgcg	ggtccgcgct	gggacgcgta	32160
cttcgcgggc	cgcggcgccc	agtggctgga	ccttcgcgacg	taccccttcc	agcgcggccg	32220
cttctggccg	gagtccttct	cgggcgcgcg	atcgggcgcg	ccggcagccg	gacagccggc	32280
cgagaccgac	gcggccttct	gggacgcctg	cgcacaggag	gacttcaccg	cattggaatc	32340
cgtactcgac	gtcgagagcg	acgcactgtc	caaggtgtcg	ccggccctga	tggactggcg	32400
cagccgccaag	gccgacgagt	cccaactggc	aggctggcgc	caccgcatcg	tctggaagcg	32460
gctcaccggc	gccgcccttg	cacaccgcaa	ggcgctcagc	ggcacctggc	tcgcggtggt	32520
ccccgagggc	ttcgccgacg	accctggggt	gaccaccacc	ctggacggcc	tcggtacca	32580
cctcgtgcat	ctggaggtcg	cggaggccga	ccgggcccgc	ctggccgacg	cgatcgcggc	32640
ccgcaccgcc	gacggcaccc	gcttcggcgg	cgtaatctcc	ctgctggccc	tgcgcgagga	32700
gctcaccggc	gcggtgcccc	aggggaccgc	cctgaccacc	accctcctcc	aggccctcgg	32760
cgacgccggc	gtcgacgcac	cgctgtggtg	cgtcaccgcg	agcgcctct	ccgccggccg	32820
caccgaccgg	ccgcaccgac	cgctccaagg	cgccgtctgg	ggcctggggc	gggtcgcggc	32880
ccttgagtac	ccgcagcgct	ggggcggcct	ggtggacctg	ccggaggagc	ccgacgagcg	32940
gtccgcggcc	ggcctcgccg	ccgtcctggc	cggtctggac	ggcgaggacc	aggctcgccgt	33000
gcgcggcacc	gcggtgctcg	cccgcgcct	ggtgccggct	cccggccgca	agccgtcccg	33060
gccctggcac	ccgtccggca	ccgtcctggt	caccggcggc	accggcgccc	tcggcgcgca	33120
cgtcgcccgc	cgcctggcca	aggacggcgc	ccagcacctc	gtcctgctca	gccgcgcgcg	33180
cccggacgct	cccgggtgcg	cggaaactgc	cgcggaactg	gacgcgttgg	gcaccgacgt	33240
cacggtcgcc	gcctgcgacg	tcgccgaccg	cgaccagctg	acggccgtcc	tggacgcgct	33300
gccgcgcgac	cggcgcgtga	ccggtgtggt	gcacaccgcg	ggcgctcctg	acgacggcgt	33360
actggaccgg	cgcacccccg	agcggttcca	ggaggtgttc	cgcgccaagg	tcacctcggc	33420
cctgctgctg	gacgagctga	cccgcgaccg	cgagctggcc	gcgttcgtcc	tcttctcttc	33480
cgctccgccc	gcggtcggca	accggggcca	ggccaactac	gccgctgcca	acgccgtcct	33540
ggacgcgctc	gccgaacagc	gccgggtgct	ggcctgccc	gccacctcg	tctcctgggg	33600
tgcctgggga	ggggcgggca	tggccgacgc	cgacggcgcg	gacgaggccg	cccggcgcg	33660
cggcgtcggc	gccatggacc	cgcacctcgc	cgtggaagcc	ctgctgcgcc	tggctgcgca	33720
gaaggagccg	accgcggtgg	tcgccgaggt	ggccttgga	cggttcgccg	gcgccttcgg	33780
cggcagccga	cccagcgccc	tgtcgcgga	gttccccggc	taccgcgagg	cgctcgccgc	33840
ccaggcggag	caggccgcgg	acggcgggcg	gctggccgcg	cgactggccg	cgctgcgcgc	33900
cgcccgccgc	ctggacaccg	ttgtggacct	ggtgcgcacc	cgcgccgcgc	aggtgctcgg	33960
ctaccccgac	accgaagcgg	tcgccgcgca	acggctcctc	cgcgacctgg	gtgtcgactc	34020
gctcggcgcc	gtcgagctgc	gcaaccaact	gagcgcgccc	accggcctga	acctgccggc	34080
gacgtggtg	ttcgaccacc	cgacccccct	ggtcctgggg	gagcacatcc	tcggcgggct	34140
cttcccggac	gagcccgccg	ggtccgacga	cgagacggag	atccggggcc	tgttggtcct	34200
cgtcccgtc	gaccaactgc	gggagatcgg	ggtcctggag	cccctgctcc	agctcgccgg	34260
acggcgcggc	cgggcgcgcg	acggcgacga	cggcgagctc	gtcgactcga	tgacagtggc	34320
agacctggtg	cgggcgcgcg	tcaacggcca	gtccgacctg	tagcgcgatt	gatggagcag	34380
acgatgaacg	cgcccagaaa	ccccgagacc	cccagaaaca	acgtagtcgc	cgactccgc	34440
gccgcggtca	aggagaccga	ccggctccgg	cggcagaacc	ggatgctggt	cgcggcgccc	34500
aagggaaccga	tcgccgtggt	cggcatggcc	tgccgcttcc	ccggcgccgt	cgactccccg	34560
gaagcgctgt	gggagatggt	cgccaccggc	accgacgtga	tctccggatt	ccccgacgac	34620
cgcggctggg	acctggaggc	gctgcgcaac	agcggcaccg	acgcccgcg	caccgacgtc	34680
agccagcgcg	gcggattcct	ggactgcata	gccgacttcg	accccggtt	cttcgggata	34740
tcaccgcgcg	aggcggtcac	catggaccgc	caacagcggc	tcctgctgac	caccgcctgg	34800
gaggcgtcg	agcgggcccg	catcgacgcc	accacgtgc	gcgccaccgc	caccggcgcg	34860
ttcatcgcca	ccaacggcca	ggactacgcc	tacctgctcg	tcctgctcct	ggacgacgcc	34920
accggcgacg	tcggcaccgg	catcgccgcg	agcgccgcct	ccggtcggct	ctcctacacc	34980
ctcggcctcg	aaggccccgc	gctcaccgtc	gacaccgcct	gctcctcgtc	gctggtcgcc	35040
ctgcacctgg	ccgtgcaggc	gctgcgcaac	ggcgagtgcg	gcatggcgct	ggccggcgcc	35100
gtcaacgtga	tggccacacc	gggtcgcgtg	gtcgagttca	gccgccaggg	cgggctggcc	35160
cgggacggcc	gctgcaaggc	gttcgcggac	gccgccgacg	gcaccggctg	gtccgagggc	35220
gccggcggtg	tgtgtctgga	acggctctcc	gacgcccgag	gcaacggcca	cccgggtgct	35280
gccgtggtcc	gcggctccgc	cgtcaaccag	gacggcgctt	ccaacggctt	caccgcccc	35340
aacggcccc	cccagcagcg	cgtcatccgc	caggccctcg	ccaacgcggg	cctggccacc	35400



ggcgacatcg	acgcgggtcga	ggcgacacggg	accggcacc	cgctcggcga	ccccatcgag	35460
gcgcagagca	tcctcgccac	ctacggccag	gaccgcgcc	accgggtgct	gctcgggtcg	35520
atcaagtcca	acatgggcca	caccagggc	gcgtccggcg	tcgccggcgt	gatcaagatg	35580
atcatggcga	tgccggacgg	cgctctcccg	cggaccctgc	acgtcgaccg	gccctccacc	35640
cacgtcgact	ggaccaccgg	cagcgtcgaa	ctctcaccg	acgcccacc	gtggcccgag	35700
actggcaggc	cgcgcgcac	cggcatctcc	tccttcggcg	tcagcggcac	caacgcccac	35760
gtcatcgtcg	aacaggcccc	cgacaccccc	gcccagggcg	ctgacgacac	tcgcgcccg	35820
accccgcgga	ccctgcggtg	gctgctctcc	gcccgcaccg	gcgcgcctct	gcgcgaccag	35880
gccaccgcgc	tgctcgacca	cctcgaccgc	cccagcggcg	accgcgggccc	caccgccttg	35940
gacaccgcgt	tctccctcgc	caccaccgcg	gcccgcctgg	aacaccgggt	cgccgtcgtc	36000
accggcaccg	acggcaccgc	cggacgggac	gcccgtaccg	cctgggtggc	gcacggcacc	36060
gcccccgacg	cccacgaagg	acacgcgcgc	ggacgcacc	gctgcgcggc	cctcttctcc	36120
ggccaggggcg	cccagcgctt	gggcatgggc	cgcgaactcc	acgcccgttt	cccgggtgttc	36180
gcacggggccc	tcgacaccgc	cgtggacctg	ctcgacgcgc	aactggggcg	caccctgcgg	36240
gaggtgatct	ggggcaccga	cgacgcgcgc	ctcaacgaga	ccggcttcac	ccagcccgc	36300
ctgttcgccc	tcgaggtcgc	cctctaccgc	ctgatcgaat	cctggggcgt	cgccccggac	36360
ttcgtcgccc	gccactccat	cggcgagatc	gcccgcgcgc	acgtcgccgg	ggtgttctcc	36420
ctggaggagc	cctgcacgct	ggtggccgc	cgcgcgggc	tgatgcaggc	gctgccgcgc	36480
ggcggggcgga	tggtcgccgt	cgaggccacc	gaggacgagg	tcagcccgt	gctcaccgac	36540
ggcgtcgcga	tcgcgcgat	caacggcccc	acctcgctcg	tcgtctccgg	cgacgagacc	36600
gccaccctcg	ccgtcgccgc	ccgactcgcc	gaacagggcc	gcccgcaccac	ccggctgcgg	36660
gtcagccacg	ccttccactc	gcccgtgatg	gacccgatgc	tcgcggagtt	ccgcgcggtc	36720
gccgagggcc	tgctctacgg	cgaaccgcag	atcccgggtg	tctccaacct	caccggcgcg	36780
gtcgcgcgac	gcaccctgct	cggcactgcc	gactactggg	tcgggcacgt	ccgcgcggcg	36840
gtccgcttcg	ccgacggcat	ccgcgccttc	accgacgcgc	gcgtcgggcg	cttccctcgaa	36900
ctcggccccg	acggcacgct	cgcgcctctg	gcccagcagt	ccgccccga	cgccgtctcc	36960
gtccccgtcc	tgcgcaagga	ccgggacgag	gagcccgcgc	cggtcgccgc	actggccccg	37020
ctgcacaccg	ccggcgctcc	ggtggactgg	acggcgcttc	acgcgggcac	cggcgcccac	37080
cgcaccgacc	tgccgacctc	cgccttccag	tacgagcgct	actggcccaa	ggccacctac	37140
cggccccgcg	acgccaccgg	cctcggcctg	accgccgcgc	accaccgcgt	gctcggcgcc	37200
gccatgtccg	tcgcggggtc	cgacgagctc	ctgctcaccg	gcaccctgtc	gctcggccacc	37260
caccctggc	tcgcggacca	cgtcgtcggc	ggcatggtct	tcttccccgg	caccggcttc	37320
ctggaactgg	cgggtccgcgc	cgcgaccag	gtcggctgcg	accgggtcga	ggaactcatg	37380
ctcgcgcgc	cgctgatcct	gcccgccacc	ggcaccgtcc	agatgcagat	cgcggtcggc	37440
gcccgcggac	acgacggcgg	ccgcgacctg	cgcttcttca	cccggccccg	ggacgaccgc	37500
gacgccgcct	gggcccagca	cgccaccggc	cggatcaccg	agggcgagcg	cgctctcgcc	37560
ctcgacacca	ccacctggcc	gccccgcgac	gcggaaccgc	tcgacatcga	cggcctctac	37620
gaccgctacc	gcgccaacgg	actcgactac	gggcccgtct	tcgcggcct	gcgcgcgcta	37680
tgggcccgcg	acaccgagat	ctacgccgag	gtcgccctgc	ccgaaggcac	cgccgacgcc	37740
gacgccttcg	gctgcaccc	ggccctcttc	gacgcgctcc	tgacagcac	cctcttcgcc	37800
tcgcgcgacg	gcgaagcgg	cagcctctcg	ccgttcgcct	ggaacggcgt	gtccctgcac	37860
gccgcggcg	cggacgcgct	gcgcgtcccg	atcaccagct	gcggccccga	cgccgtggag	37920
atcaccgcgc	tcgaccgca	gggcgccccc	gtcgtctccg	tcgaatcgct	gacgtgcgc	37980
ccgcgcggcc	ccgacgcgg	caccgcccac	caccgtgccg	acgcgggctc	cctcttcgcg	38040
atggactgga	cgccccgcac	cgtccacgcc	ccggccacc	ccgccacctg	ggcgtctctc	38100
ggtaccgacc	cgatcgccct	gaccgaggcg	ctcaccgcgc	ccggccccga	caccgtcacg	38160
ggactccgcg	acgggtgtcga	cgcctctggc	gaactcaccg	ccggcgacga	ccggccgggtg	38220
cccagcgtgg	tcgcgggtacc	gctgcgcggc	gccaccgacc	acgggcccgc	cgggtgccac	38280
gacctgacc	gcaccgtcct	ggcctgtctc	cagggaatggc	tgcccgagga	gcgcttcgcc	38340
cgctccccgg	tgctgctcgt	caccgcggc	gcggtcgccg	acggcgagcg	cggcccgcctc	38400
gacctggccg	ccgccccggg	ctggggcctg	gtgcgtcccg	cccagtcgga	gaaccccggc	38460
cgactgctgc	tcgtcgacct	cgacgacacc	gccgagtcgc	ccgcccaact	cccgttgctg	38520
ccggccctcc	tggaacgcga	cgagccgcag	gccgtggctc	gcgagggcac	cgtcggggtc	38580
ggccggctcg	cccgcctgga	ctccggccgc	ggcctcgctc	cgcgcgcccg	caccccttg	38640
cgcctgggca	gcggggccaa	gggcagcctc	gacggcctcg	ccctgctgcc	ccaccccgag	38700
gcccgcagac	ccctcaccgg	ccacgaggtc	cgcgtcggca	tcgcgcgcgc	gggcctgaac	38760
ttccgtgacg	tgctcaacgc	gttggggatg	tatcccgggg	atgcggggct	gttcgggttcg	38820
gaggcgggcg	gtgtgggtcgt	cgaggtcgga	ccggaggtca	ccggcctggc	acccggcgac	38880
cgggtcatgg	gcattgctctt	cggcggtctc	ggaccgctcg	gcacgcgcga	cgcccggtcg	38940
ctcacccccc	tcccggccga	ctggctcgtg	gagacgggtg	cgctgggtgcc	gttgggtgttc	39000

ctcaccgcgt	actacgccct	gaaggagttg	ggtggtctgc	gggcggggga	gaagggtgctg	39060
gtgcatgccg	gtgccggtgg	tgtcgytatg	gcggcgatcc	agatcgcccc	gcatgtcggt	39120
gccgaggtgt	tgcgcacggc	cagtgaagggc	aagtgggacg	tgctgcgctc	cctggggcgtg	39180
gccgacgacc	acatcgccctc	ctcccgcacc	ctcgacttcg	aggcggcctt	cgccgaagtc	39240
gccggcgacc	gcggcctgga	cgctcgtactg	aacgcgctgt	ccggcgagtt	cgtcgacgcc	39300
togatgcggc	tgctcggcga	cggcgggccg	tccctggaga	tgggcaagac	cgacatccgc	39360
ccgcgcgact	ccgttcccga	cggcctctcc	taccactcct	tcgacctcgg	catggtcgat	39420
ccggaacaca	tccagcggat	gctgctcgac	ctcgtcgagc	tgttcgaccg	cggcgcgctg	39480
gccgcgttgc	cggtcgcgag	ctgggacgtg	cgccgcgcgc	gcgaggcggt	ccgcttcgat	39540
agcctggccc	agcacatcgg	caagatcgtg	ctcaccgtgc	cgcaaccctt	cgaccccgac	39600
ggcaccgtgc	tcctcaccgg	cggcaccggc	ggcctggccg	gcctgctcgc	ccgccacctg	39660
gtcaccgagc	acggcgcccg	ccacctgctg	ctggccggcc	ggcgcgggcc	cgacgcgccc	39720
ggcgccgccc	cactccacgc	cgaactgacc	gccctgggcg	ccgaggtcac	cgtcgcgcgc	39780
tgcgacgtcg	ccgaccgcac	cgcgctcgcc	gcgctgctcg	ccaccgtgcc	cgccgaacac	39840
cccctcaccg	cggctcgtgca	caccgcccgc	gtcctggacg	acggcaccct	caccgcccctg	39900
aaccccgacc	gcctcgccac	cgctcctacg	cccaagggtg	acgcccgcctg	gcacctgcac	39960
gacctcacc	gccacctcga	cctggccgcg	tctgtgctct	actcctccac	cgccggcgctc	40020
atgggcggac	cgggccaggc	caactacgcg	gccggcaaca	ccttcctcga	cgcgctcgcc	40080
gcccaccgac	acgccctcgg	cctgcccgcg	acctcgctgg	cctggggcgc	ctgggagcag	40140
ggcgccggca	tgaccggcgc	actgaccgac	cacgacctgc	gccgggtcag	cgacgcgcgc	40200
ggccaaccgc	tgtcaccgc	cgaacgcggc	ctcgccctct	acgacgccgc	caccgcccgc	40260
gacgaacccc	tgatcgtccc	gctcggcctc	accggcggtg	cgctgcccgc	cggggcgggc	40320
gtccccgcg	ccgcgcgcg	cctggtccgc	accgcggggc	gccggggccag	ggccggcacc	40380
gccggcgtct	cccgcgcgg	cctcgccgaa	cgctcgcgcg	ccctgcccga	ggaggagcgc	40440
accccttcc	tcgtcgagct	ggtgcgcacc	gaggccgcca	ccgtcctcgg	ccacggctcc	40500
accgaccggg	tggacgcccg	ccgcgagttc	cgccaactcg	gcttcgactc	gctgaccgcc	40560
atcgaactgc	gcaaccgact	cggcaaggcc	accggcctca	ccctgcccgc	caccctcatc	40620
ttcgactacc	cgacccccga	ccgcctcgcc	gtccacctcc	acgacgaact	cctcggcgcg	40680
gacgccccgg	tgaccgtcac	cgcgcgcgca	caggccgcgg	acccggagca	cgacccggtc	40740
gtcatcgtcg	gcatgagctg	ccgcttcccc	ggcggcgtca	gctccccga	ggagctgtgg	40800
gacctggtgg	catccggcac	cgacgcgatc	accggcttcc	ccgccgaccg	cgcattgggac	40860
cgccaccggc	agctcgccgg	cgcccccggc	gcccgcaccg	gccagggcgg	attcctccgc	40920
gacatcgccg	acttcgacgc	cgccttcttc	ggcatctcgc	cgcgcgaggg	cctggccatg	40980
gacccgcagc	agcgcatcct	cctcgaagtc	gcctgggagg	ccgccgagcg	cgccggcatc	41040
gacccgcaga	ccctgcgcgg	cagcgacacc	ggcgtgttca	tgggcgctcag	cggccaggac	41100
tacgccggcc	tcgtgatgcg	ctcccgcgac	gacatcgccg	gccacgccac	caccggcctc	41160
gccgtcagcg	tcgtctccgg	ccgcctcgcc	tacgcgctcg	gcctggaggg	cccggccctg	41220
tccgtggaca	ccgcctgctc	ctcctccctg	gtgtcgctgc	acctggccgc	ccaggcgctg	41280
cgcgcggggg	agtgcacat	ggccctggcc	ggcggcgtca	ccgtcatgac	caccgcccgc	41340
aaattcaccg	gcttctcccg	gatgggcggc	ctcgcccagg	acggccgctg	caaggcgctt	41400
tccgactccg	ccgacggcac	cggctggtcc	gagggcgccg	ccgtcctggt	cctggaaacgc	41460
ctctccgacg	cccggcgcg	cggccaccgc	gtactggccg	tggtgcgcg	ctcggcggtc	41520
aaccaggacg	gtgcgtccaa	cggctctgac	gcgcccacg	gtcccgccca	gcagcgcgctc	41580
atccggcagg	ccctggccaa	cgggggcctg	accccgctcg	acgtggacgc	cgtcgaggcg	41640
cacggcaccg	gcaccccgct	cggcgacccc	atcgaggccc	aggccctgat	cgccgcctac	41700
ggcaccgacc	gcgaccccg	acaccgcgtg	ctgctcggtc	cgggtgaagtc	caacatcggc	41760
cacacccagt	ccgcggccgg	cgcggccggg	ctggtcaaga	tggatcatgg	catgcgccac	41820
ggcatcctgc	cgcagacct	gcacctcacc	gaaccgtcct	cgcacgtgga	ctggtcggcg	41880
ggcacggtgc	ggctgctcac	cgagcggacc	gcctggccgc	ggacggatcg	tccgcgtcgg	41940
gccgggggtc	cctcgttcgg	catcagcggc	accaacgccc	acgtcatcct	ggaacagccg	42000
cccgcgcgac	ccaccccgc	cgcgcgaccc	ggccggcccc	caccacccgt	ggtggccctg	42060
cccgtctccg	cgcagacccc	ggccgccttc	gacgcccaac	tggaccgggt	gcgcaccgcc	42120
gccgccctgg	cgcgcctcga	caccgcccac	acctcgcca	ccggccgctc	gctcttcgaa	42180
caccgcgcgc	tcctgctcgc	caccgtcggc	gacccggcga	ccggcgcccc	cgacctgcc	42240
gaggtcgcca	ggggagcggc	gacgccgcac	cgcaccgcgt	tcctcttctc	ggggcagggg	42300
gctcagcggg	cggggatggg	gcgtgaactg	catgctgctt	tcccggtgtt	cgcggcgggc	42360
ttcgacgagg	tgggtggctgt	gttggatgcg	gagctgggtt	ctgatgctga	tgggggtgtg	42420
tcgctgcggg	aggtgatgtg	gggcgggggg	tcggagtgtg	tggatcgaac	gcgtttcacg	42480
cagccggcgt	tgttcgcggg	ggaggtggcg	ttgttcggtt	tggtgccctc	gtgggggggtg	42540
gggcctgagt	tcgtggcggg	gcattcgggtg	ggtgagattg	cggcggcgca	tgtggccggg	42600

gtgttctcgt	tgggtgatgc	gtgtcgtttg	gtgggtggcgc	gggcttcggt	gatggatgcg	42660
ttgccgggtg	gtggcgatgat	ggttgcgggtg	gagggcgccg	agggcgaggt	ggtgccgctg	42720
ttgggtcgatg	gggtggcgat	cgccgcgggtc	aacggggccg	tctcgggtggt	ggtctccggt	42780
gtggaggcgg	cggttgggca	ggtcgtggat	cagttgggtg	agcggggccg	gcgggtccgt	42840
cggttggcgg	tcagtcatgc	tttccactcg	cggttgatgg	atccgatggt	ggatgccttc	42900
cggggcggtc	ccgagggcct	ggagtaccac	cagccgcgca	tccccgtggt	gtccaacgtg	42960
acgggcgagg	tggccgcggc	ggaggagctg	tgcgcggccg	actactgggt	gcggcacgtc	43020
cgggcgaccg	tgcgggttcgc	cgacggcgctc	cgcaccctgg	ccgagcgccg	cggcaccgcc	43080
ttcctggaga	tcggccccga	cggcgactctg	tccgcgctcg	cccgcggcgt	cctgccccgc	43140
gaggcgctcg	tgacgcccac	cctccgcaag	gaccgcgacg	aggagagcgc	cctgctcgcc	43200
ggactggccc	ggctgcacgt	cgccggcggtg	accgtcgact	ggagcgccgc	cctgaccggc	43260
accggcgccg	gcggcaccga	cctgcccagc	tacgccttcc	aacgcgagcg	gtactggccg	43320
gagttggccg	ccgaacccgc	gggcggcgccg	gcggatgccc	cggacgcgga	gttctggggc	43380
gcgggtggagc	gcgcggacgc	caccgcgctc	gccgcccacc	tggacatcga	cggcgaccag	43440
ctcggcgccg	tgctgcccgc	actgtccgcc	tggcgcaccc	ggcgccgcac	cacatcggcc	43500
accaacgccc	tgcggcaccg	ggagagttgg	gaaccgctgt	cgctcgccgg	cacgccgcac	43560
accggcgggc	tcctgggtgct	ggtgcccgcgc	gccgcgacca	ccgacccctg	ggtcgccgac	43620
gtcgtcgccg	ccctcgcccc	ggacgcccgc	cgggtggacg	tcccggccga	cggcaccgac	43680
cgggcccgcgc	tcgcccgcct	gttcaccgaa	gcggccgacg	acaccgcccc	gaccgcccgtg	43740
gtctccctgc	tcgcgctcga	cgagaccagc	ggcgacgacg	cggtagccggc	cggcaccacc	43800
gccaccgccc	cgtcgtccca	ggccctcgcc	gacaccggcg	ccccggcccc	gctgtggggc	43860
ctgacccggg	gcgcggctcg	cgcgctcccc	gacgagcagc	cgaccgcccc	cgcccaggcc	43920
gccgtctggg	gcctcgcccg	gatcgccgcc	ctcgaactcc	cgcgccactg	gggcggactg	43980
gtcgacctgc	ccgccgacct	cgacgagcgc	accgcacgcc	gactgcccgc	cgactggccc	44040
gacgcggcg	acgaggacca	gctcgcgctg	cgggccaccg	gcgcctacgg	ccgccggatc	44100
accccgggcg	ccgccccga	cgacgcccc	ggcaccgggt	ggcagccgac	cggcaccgtc	44160
ctgatcaccg	gcggcaccgg	cgcgctcgcc	cggcacaccg	cccgttggct	cgccgcccac	44220
ggcgccgagc	acctgctgct	gctcagccgc	agcgcccccg	acgcgcccgg	cgcgccgaa	44280
ctcaccaccg	aactcaccgc	cctcggcgcc	cgcgtcacc	tctgtggcctg	cgacgcccgc	44340
gaccgggagc	agctgaccag	ggtcctcgcc	gaggtaccgc	gggactgccc	gctgaccggc	44400
gtgggtgcaca	ccgccggagt	gctcgacgac	ggcgtgctca	ccggcctcac	cccggaccgg	44460
ttcgccacgg	tcttcgcgc	caaggtggcc	tccgccgtgc	tcttcgacga	gctgaccggg	44520
gacaccgacc	tggccgtctt	cgcgctgttc	tgcctcgctg	cgggcgctgt	cggcaacccc	44580
ggccaggccg	gctacgccgc	cgccaacgcg	gtcctcgacg	ccctcgccgc	ccgccgcggg	44640
gcccaggggc	tggccggcac	ctcgatcgcc	tgggggtgct	gggcccggca	cggcatggcg	44700
gcccgtcaca	cccgccccgg	cgcggaaccc	gtcggcctgc	tcgaccccga	cctcgccgta	44760
ccggcccttg	ccgcgcgggt	gacggagccc	cagcccaccc	tcttctcgc	cgacctccag	44820
cagccgcggc	tgttgaatc	cctgctggcg	ctgcgcccc	gcccgtcctt	gagccggctg	44880
cccgcgcgcc	gcaccgcggc	ccgcgcgggtc	caggaagcgg	accgcgcgcg	agccggcgcc	44940
gcggccgacc	tgcgcgacca	actcgccggc	accgcacccg	ccgaccgcca	cgccgtcctc	45000
ctccgcctgg	tgccgaccac	ggccgcgcgg	gtcctcgccg	acaccggtgc	cgacgccatc	45060
cgggcccga	agcccttcgc	cgacctcgcc	ttcgactcgc	tcaccgcggg	ggaactgagc	45120
agcgtctcgc	ccgcgcgcac	cggcctggcc	ctcccgccca	gcctcgtctt	cgaccacccc	45180
tccccgcggg	cgctcgccga	ccacctgcgc	gccgaactca	ccggcgaccg	gccggaatcc	45240
gccccgcggg	ccccgcgggc	accggtcccc	gccgcggacg	acgatccgat	cgtcgtcgtc	45300
ggcatggcct	gccgcttccc	cggcggcgctc	accacccccg	aggagtctctg	gcagctgctc	45360
gccgagggcc	gggacggcat	cgacgcgttc	cccaccgacc	gcggctggga	cctcgacgtg	45420
ctcggccggc	gacggccagg	gccacagcgc	ccaccgaggt	cggcggtctc	ctcttacgac	45480
gcggccgcct	tcgaccccgg	cttcttcgac	atctccccgc	gcgaggcgct	cgccatggac	45540
ccgcagcagc	ggctgctgct	ggagaccgcc	tgggaggccg	tcgaacgcac	cggcaccgac	45600
ccgacccggc	tgcgcggcag	ccgcaccggc	gtgttcgctg	gcaccaacgg	ccaggactac	45660
gccggcctcg	tcctgcgcgc	ccaggaggac	gtcgaggggc	acgccggcac	cggactggcc	45720
gccagcgtga	tctccggccg	cctcgccctac	gccttcgggt	tcgagggccc	cgcgctcacc	45780
gtcgacaccg	cctgctcctc	ctcgctcgctc	gccctgcaact	gggcgctcca	ggcgctgcgc	45840
gcgggggagt	gctccctggc	cctggccggc	ggcgctcaccg	tcatgacgac	ctcgacgagc	45900
ttcgccggct	tcaccgggca	gggcggcctg	gcgcgggacg	ggcactgcaa	ggcgttctcc	45960
gactccgcgc	acggcaccgg	ctggctcgag	ggtgtggggc	tcttcgctcg	cgaacgcgcg	46020
tccgacgcgc	tcgcgaacgg	ccatgagatc	ctggccgtgg	tgcgcggctc	ggcggtcaac	46080
caggacgggtg	cgtccaacgg	tctgacggcg	cccaacggtc	ccgcccagca	gcgcgtcatc	46140
cggcaggccc	tggccaacgc	gggcctggcc	cccggcgacg	tggacggcgt	cgaggcccac	46200

ggcaccggca	ccgtcctcgg	cgaccccatc	gaggcccagg	cgctgctcgc	cacctacggc	46260
caggaccggc	ccgccgaccg	gccgttgtgg	ctcggtcctg	tgaagtccaa	catcggccac	46320
acccaggccg	ccgcgggcgc	cgccggcctg	atgaagatgg	tgctggccct	ccaacacggc	46380
acgctgccgc	gcaccttgca	cgtcaccgag	ccctcgaccc	gggtcgactg	gtcggccggc	46440
gcggtgcggc	tgctcaccga	gcggaccgtc	tggccgcgga	cggatcgtec	gcgtcgggcc	46500
ggggtctcct	cgttcggcat	cagcggcacc	aacgccacgc	tcatacctgga	acagccggcc	46560
ggcgagccca	ccccacggc	ccctgccgac	cgccccaccc	ggacgccgcg	cgtcctccca	46620
tgggtcgtct	cggcccgatc	ggccaccgcg	ctcgacgcgc	agctcgcgcg	actcggggcg	46680
ttcgccgcgc	agcgcgcgga	cctgccgcgc	gccgacgtcg	cccactcgct	cgtcaccagc	46740
cgcgccacct	tcgaacaccg	ggcgggtcctg	ctggccgcgc	ccgacggcat	caccgcggcc	46800
gcccgcgcgc	aggcccgcga	acgcagcacc	gcgttcctct	tctcggggca	gggtgctcag	46860
cggtcgggga	tggggcgtga	actgcatgct	gctttcccg	tggtcgcggc	ggcgttcgac	46920
gaggtggtgg	cgggtgttga	tgcggagttg	gcgacggggt	ccggtggggg	tgtgtcgctg	46980
cgggaggtga	tgtggggcgg	ggggtcggag	ttgttggatc	ggacgcgttt	cacgcagccg	47040
gcgttggttcg	cgggtggaggt	ggcgttggtc	cgtttgggtg	cctcgtgggg	gggtggggcct	47100
gagttcgtgg	cggggcattc	ggtgggtgag	attgcggcgg	cgtatgtggc	cgggggtgttc	47160
tcgttggtgg	atgcgtgtcg	tttgggtggg	gcgcgggctt	cgttgatgga	tgcttgccg	47220
gtgggtggcg	tgatggttgc	ggtggaggcg	gccgaggcgg	aggtggtgcc	gctgttggtc	47280
gatggggtgg	cgatcgccgc	ggtcaacggg	ccggtttcgg	tggtggtctc	cgggtgtggag	47340
gcggccgttg	ggcaggctcg	ggatcagttg	gtggagcggg	gccggcgggt	ccgtcgggtg	47400
gcggtcagtc	atgctttcca	ctcgccgttg	atggatccga	tggtggatgc	cttcggggcc	47460
gtcgccaggg	gcctggagta	ccaccagccg	cgcatacccg	tggtgtccaa	cgtgacgggc	47520
gaggtggccg	cggcgaggga	gctgtgcgcg	gccgactact	gggtgcccga	cgtccgggcg	47580
accgtgcggt	tcgccgacgg	cgtccgcacc	ctggccgagc	gcggcggcac	cgccttcctg	47640
gagatcggcc	ccgacggcgt	actgtccgcc	ctggccgcgg	cctgcctggt	cgacacggac	47700
gccgaagtgg	tgcccgcgct	gcgcaagggg	cgccccgagg	agcacaccgc	cctcaccgcc	47760
gccgcccaac	tcacagtggc	cggcgtggac	atcgactgga	ccgcggtcct	ggccggcacc	47820
ggcgggcccgc	ggatcgccct	gcccacctat	gccttcacgc	gcgagcggta	ctggccctcg	47880
ctcgccgcac	aggcccccg	cgacgcgcgg	gggtcggcc	tggaaagccg	gcggcaccgc	47940
ctgctcgggg	ccgcgaccac	cgtcgccgga	tccgcggaga	tctgtctcac	cggccgcctg	48000
tccaccaccg	cccagccgtg	gctcgccggt	tacgaggcgg	acggccgcac	cgtcctgcgc	48060
gccgcggtcc	tcgccgaact	cgcgcgtccg	gccggcgacc	aggccgactg	cccagaccgtg	48120
gcggaactga	ccgtcgccgc	accgctcgte	ctcaccggcg	cggcgggcca	gcgcctccag	48180
gtccgggtgg	ccgcccccca	cgacaccggc	cggcgcgcg	tgctcgtgca	cgcccgaccc	48240
gacgactccc	ccgacagccc	ctggacgctg	cacgccaccg	cggtectcac	ccacgacacc	48300
ccgcagcccc	cggcgccgga	caccggctgg	ccgccggagc	gcgcgctgcc	gctcgacgcc	48360
ctgcccaccg	ccaccggccc	ggcccggatc	gcggcgccct	ggcagtgggg	cgacgaactc	48420
tgccgcgaga	tcgaactccc	cgaacccggc	ccggcgaggc	gggcattcgc	cctgcacccg	48480
gcgctgctgg	acaccgcggg	ccgcgcgggc	ggcctgctgg	acggcgacgc	caccctggac	48540
gccctcgggt	ggcggggcct	cgccttcgac	gccgcgtccg	ccaccgcctt	gcgggtccgc	48600
ctcaccgccg	acggcacgga	cacctgggct	ctggaggcca	ccgaccgcga	gggcgtcccg	48660
gtcgtctccg	tcaccgggct	caccctgggc	acgcccaccg	tcgaccggtc	gggggcccggg	48720
gcggccgatg	acggcgcgac	cctgctcgac	ctggagtggg	tgcccgcgcg	gcaggccgcg	48780
cccaccggcg	gcgaccacct	cccgtacgcc	gtgctcggcg	atcaactcgc	ggagctggac	48840
gggcagttga	ggatcgcggg	cgacggggcc	gggcgcgtcg	catcgctggc	cgcgctgctg	48900
gacggcggtg	cgcgcgtgcc	ccggctcgte	ctcgcgccgg	tgctgggcgt	gccgaccggg	48960
gaaggcgacc	tgcccgcgcg	ggtgcgcggc	accaccacgg	cgggtgctgga	gctgctgcag	49020
cgtgggaccg	ccgacgcccc	caccgcccgc	agccacctgg	tgatcgteac	ccgcggcgcc	49080
gtcgccgccc	gggcggagga	cgtgcacgac	ctggcgggcg	ccccggctctg	gggcctggtc	49140
cgtcgggcac	agtccgaaca	ccccggcagc	ttcctgctgc	tcgacctcga	ccccgcgat	49200
cccgcgggag	cctcccgcgc	cgcgcgcggc	gccaccctgg	cggccctgct	cgacgcgggc	49260
gagaccacag	ccgcggtgcg	cgcgcgacac	ctcaccgtcg	cccggctgac	ccgggcccgc	49320
gacggacccc	aggccaccgc	cggacacccc	gtgcgggact	gggaccgcga	cggcaccgtc	49380
ctgatcaccg	gcggcaccgg	cggcctgggc	ggcctcctgg	cccgccacct	ggtcaccgga	49440
cacggcatca	agcacctgct	gctcgccggg	cgcgcggggc	cggacgcccc	cggcgcgcg	49500
gccctgcgcg	acgaactggc	cgccttcggc	gccgaggtga	ccgtcgccgc	ctgcgacgtg	49560
gccgaccgtg	ccgcactgga	ccgactcctc	gcgcaactgc	cgcggagga	cccgtgacc	49620
gccgtcgtgc	acaccgcccg	cgtcctcgac	gacgccaccg	tcggcaccct	gacgcccag	49680
cggctggaca	ccgtcctgcg	cgcgaaggcg	gacgcgcct	ggcacctgca	cgacgccacc	49740
cgcgaccgcg	acctggcagg	gttcgtgctg	tactcctcgg	tcgccgggtg	caccggcggc	49800

cccggccagg	gcaactacgc	cgccggcaac	acgttctctg	acgcgctcgc	cgcgcaccgc	49860
gccgcccagg	gcctgcccgg	actgtcgctg	gcctggggac	cgtgggggca	ggacgcgggc	49920
atgaccggca	ccctcgggcg	cgccgacctg	gcccgcctgg	agcgctccgg	catgccgcgg	49980
ctcaccgccg	aacagggcct	ggccctgttc	gacgcgcgcg	gcgcccgcgg	cgacgggttc	50040
gcggtggcgg	tgcggctcgc	ccgtggcgcc	gccgcaccgg	gcgcccagca	ggtccccgcg	50100
gtgctgcgtg	ccctgggtgc	cggccggcgc	cgcacggcgg	ccgcgccggg	gcacgcgggt	50160
gtactggccc	gcccggctgg	cgccctggac	gcccagcagc	ggcatcaggc	gctgctcgac	50220
ctgggtccgca	ccgagacggc	cgcgggtgct	ggccactccg	gggcggacgc	cgtcccggcc	50280
gagcgggact	tcaaccgggt	gggcttcgac	tcgctgatgg	cggtcgaact	gcggacgcgg	50340
ctggccaccg	ccaccggagc	ccggctgccg	gccacgctcg	tcttcgacca	cccagcgcgg	50400
gacgcggctg	cccggcacct	cgcgtcgacg	ctgcccggtg	ggaccgcggc	cggtcgggac	50460
cgttccccgc	tggccgaact	cgaccggatc	gccgcggagt	tgtcgccgga	gggcgcggac	50520
gacgccaccc	gacagggcgt	cgtcgggcgg	ctgcggcacc	tgtggcgca	gtgggacggc	50580
acccgacagg	acggcgggtg	gacgaccgtc	gacgaccgca	tgaagcggc	gagcgcgcaa	50640
gaggtcctcg	ccttcacatg	ccacgagctc	ggccggcagg	cggactcctg	accgccccca	50700
ctcccgtcgc	tcgcgcgcac	cacatctgag	gaagggttca	cggaccatgc	cggacgaaaa	50760
gaagctcgtc	gactatctga	agtgggtcac	gaaggacctc	caccagaccc	gccagcgcct	50820
tcaggagggtg	gaggcggggc	gccacgaacc	cgtggcgatc	gtcggcatgg	cctgccgcctt	50880
ccccggcggt	gtgcgctccc	cggaggacct	gtgggagctg	ctgtccgcgg	gccgggacgg	50940
catcgggccg	ttccccgcgg	accgcggctg	ggacctggcg	gcgctggccg	gcgacggggc	51000
cggtcgcagc	gccacccagg	aaggcgggtt	cctgcccagc	gcggccgcct	tcgacccggg	51060
cttcttcgac	atctccccgc	gcgaggcgct	cgccatggac	ccgcagcagc	ggctgctgct	51120
ggagaccgcc	tgggaggccg	tcgaacgctc	cggcatcgac	ccggccgggg	tgcgcggcag	51180
ccgacccggc	gttttcgctg	gcaccaacgg	ccaggactac	gcgcacctgg	tctcgccgcg	51240
gcaggacgac	tggggcggct	acgcgggcaa	ggccttgccc	gccagcgtgc	tctccggccg	51300
actggccttc	gcgctcgccc	tggaaaggccc	ggcgcgtcac	ctcgacaccc	cctgctctct	51360
gtcactgggtg	accctgcacc	tggccgcaca	ggcgcgtcgc	gccggcgaat	gcggcctcgc	51420
cctggccgggt	ggcgtcacgg	tcatgacgac	ctcgtcgagc	ttcgccgggt	tcagcctcca	51480
gggcgccctg	gcgcgggacg	gccgctgcaa	ggcgttcgcc	gaggcgcccg	acggcacccg	51540
ctgggtccgag	ggcatcggcc	tgtctctcgt	cgagcggctc	tccgacgcgc	agcgcaacgg	51600
ccaccgggtg	ctcgccgtgc	tgcgcggctc	cgccgtcaac	caggacggcg	cgtccaacgg	51660
cctcagcgcg	cccaacggct	cgtcccagca	gcgggtcatc	cgccaggcgc	tggccggcgc	51720
cggactcgtc	cccggcgacg	tggacgcggg	cgaggcgcac	ggcaccggca	cccggctcgg	51780
cgaccccatc	gaggccgggt	cgctgctcgc	cacctacggc	caggaccggc	ccgcccagcc	51840
gccgttgttg	ctcggctcgg	tgaagtccaa	cctcggccac	accagggccg	ccgcgggcgt	51900
cgccggcgct	atcaagatgg	tgttgccctt	gcggcatggc	gtcctcccgc	agacctgca	51960
cgtggacgcg	ccctcctcgc	acgtcgactg	ggagagcggc	gcggtgcggc	tgtcaccgcg	52020
accgctcgcc	tggctcgagg	gcgacgaccg	ggtgcgccgg	gccggcgctc	cgtcgttcgg	52080
catcagcggc	accaacgccc	acgtcatcct	cgaacaagcc	cccgatcagc	cggaaaccgac	52140
cgcggaagag	acggctgcgg	cggcgcccgg	cggcaccgcc	gaggagcggg	ccgcgcgtcc	52200
cgtcgccccg	cgcgcctgtc	cgtggccggg	cgcggcacgc	accgcggcgg	ccctcgacgc	52260
ccaactggct	cgggtccgcg	cgctgaccac	cgcgcaccgg	cgcaccgcgg	ccgacgtcgg	52320
tcacgcgctg	gccaccgccc	gtaccgccct	cgagcaccgg	gcgctgctgg	tccacgaggg	52380
cggcgcgcgt	accgaggtgg	cgcgcggcgc	cgtccccacc	ggtgaccggg	gcgggctggc	52440
cgtgctgttc	tccggacagg	gctcccaacg	gccgggcatt	gggcgcgaac	tccacgcccg	52500
ctaccgggtc	ttcgccgcgg	ccttcgacga	gaccgtcgcc	ctgctcgacg	cccggctcgg	52560
cacgtcgttg	cgcgacatcg	tctgggacca	ggaccgcacc	cggctcgacg	acaccgcgca	52620
caccagcccc	gcgctgttcg	ccgtcgaggt	cgcgctgtac	cgcctgctgg	cctcctgggg	52680
catccggccc	gaccacgtca	ccggacactc	catcggcgag	atcaccgcgg	cgcacgtcgc	52740
cgggtgtgctg	accctcgcgg	acgcctgcac	cctggtggcc	gcccgcgcca	ccgccatgag	52800
cgaactgccg	cccggcgggc	ccatgggtgg	gctggaggcc	accgaggacg	aggtgcgtcc	52860
gctgctcacc	gacgacctcg	cgatcgccgc	ggtcaacgcc	ccccggtccg	tggctgctcg	52920
cggcgccgag	gacgcgcgcc	tcgccgtccg	ccggcacttc	gacgacctgg	gccgcgggac	52980
cacccggttc	ccggtcagcc	acgccttcca	ctcgccgctc	atggaccgca	tgtcgcagcg	53040
cttcgggacg	gccctcgcgc	cgctgacctt	cgcgcagccg	gagatcccgg	tcgtctccaa	53100
cctcaccggc	ctcccggcca	ccgccgagga	actcgccacc	ccgcactact	gggtgtgcca	53160
cgtccggcag	gccgtccgct	tcggcgacgg	cgtgcgcgcc	ctcgccgacc	gcggcgtgcg	53220
gaccttcttc	gaactcggcc	cggacggcgt	gctgtccgcc	ctggtccggg	agaacctccc	53280
cgagccgggc	ctggctcgccg	tgcctgtgct	gcgcaaggag	cggcccggag	agaccaccgt	53340
gctggccgcc	ctgggaaccc	tgtgggcgca	cggcgccggac	gtggactggg	acgcgggtgt	53400

cgccggcacc	cgcacccccg	aggccgaccc	cgtcgagctg	ccgacgtacg	ccttccaacg	53460
cgcccgtac	tggcccaccc	tcggcgcccc	ccacggcgac	ccggccgacc	tcgggcagac	53520
cgccggccgc	cacccgctgc	tggggcggcg	cgtcaccctc	gccgacgccg	acgagaccgt	53580
gctcaccggc	cgcctcggcg	tgccctccca	cccctggctc	ggcgaccacc	gcagcgacgg	53640
ccggatcacc	gtccccggcg	tcgccttcgc	cgaactcgcc	gtccgcggcg	gcgacctgag	53700
cggcaccgcc	cacctggcg	ggctcgacct	gccggcgccg	ctcaccctcg	gcgacggcga	53760
caccgtcacc	ctccaggctc	gggtcgggcg	ccccgacccc	gcggggccacc	ggccgctgac	53820
cgtccacgcc	cgcctcgcag	ccaccgagga	cgcacctggg	accacctgcy	cgaccggtct	53880
gctcgccccg	gacgcccccg	aagcgccccg	ggatccgata	ggccccggcg	acgcccgggtg	53940
gccgcgcggg	gacgcccccg	cgggtgcccgt	gcgcgacctc	gacgcggccg	ccaccgcgcg	54000
aggccgccac	tacggcccc	atttccaggg	cctgaccggg	ctctggcggc	gcgacggcga	54060
ggtcttcgcc	gaggtggccc	tgccaccgc	caccgcgcgc	gaccgcgcct	tcggcatcca	54120
ccccgcgctg	ctggccaccg	cgtccgcgc	caccgcgcga	ctggacgacg	accacaccgc	54180
cggccacacc	cccgaaccga	ccggcatcac	cggactcgcc	ctgcacgcca	ccggggccac	54240
cgcactgcgg	gtccggctca	ccgcgaccgg	gcccgcaccc	gtggccctcg	ccgcgcggga	54300
cgccacgggg	ggcgcggtcc	tgaccgccga	caccgtcacc	ctcggtctcc	cgcaggaccg	54360
cccggctccc	gcaccggccg	gccacaccgg	gcagggcggc	ctgttccacc	tcgactgggt	54420
gccggtcgac	cccggcagcc	gagccaccgg	caccgcgtgg	gccgtcgctg	gcgacgacga	54480
actcgacctc	ggctacgccc	tgcaccgcgc	cgacgagacg	gtcagtgcct	acgcggcgctc	54540
gctgggcgga	gccatcgggg	acagcggtct	ggcgcccgcg	gtcttctctg	tccccgtcgt	54600
cggcgggccg	gacgcggggc	ccgacgcggt	gcacgcgcgc	accgcccgcg	ccctgggggt	54660
gctccaggag	tggctgaacg	agccgcgggt	ggccggcgcc	cgcctgggtct	tcgtcaccgc	54720
cggcgccgtc	gcggtgcccc	gcgagaccgt	caccgacccg	gccggcgccg	ccgtctgggg	54780
cctgctgcgc	tccgcccaga	ccgagaacct	gggcagctcg	ctgctggctg	acctcgacga	54840
cgcgttccgg	tccgcgggga	tgctgcccga	cgctctcacc	ctcgacgaac	agcagctcgt	54900
cgtccgcgac	cacgcgggtc	gcgcgcggcg	cctggccccg	ctgcccggagc	cggccgcccg	54960
caccgcgcgc	gcccgcgcct	gggacccgga	cggcaccgtc	ctgatcaccg	gcggcaccgg	55020
cggcctgggg	gcccgcgctg	cccgccacct	ggtcaccgtc	cgcggcgccc	gccacctgct	55080
gctcgccggc	cgcgcggggc	ccgaggcgcc	gggcgcggcg	gaactgggtg	cggagctgac	55140
cgcacagggg	gcggacgtgc	gggtggccgc	ctgcgacgtc	ggcgaccgca	ccgcccctga	55200
cgcgctcctc	gccacgggtc	ccgcggcgca	cccgtgacc	gccgtcgctg	acaccgcggg	55260
cgtcctggac	gacgcccctg	tcggctcgct	cacccccgac	caactggcca	ccgtgctaag	55320
gccaaggcc	gacgcccctc	ggcatctgca	cgacgccacc	cgcggccctg	acctggccgg	55380
cttcgtcctg	tactcctcgg	tctccggcgt	cctgggcagc	cccggccagg	gcaactacgc	55440
cgcgcgcaac	gcctacctcg	acgcgctcgc	ccggcaccgc	gccgaccagg	gcctcccggc	55500
gctctccctc	gcctggggcc	cctgggggtc	gggcagcggc	atgaccgcgt	cggtcagcga	55560
cgcgcgacct	gagcggtatg	cgcgcggcg	cctgccgcgg	ctgaccgtcg	aggacggcct	55620
ggccctgttc	gacgcgcggc	tcggccgccc	cgagccggcc	ctggtgccca	gccgcatcaa	55680
cgtcgccggc	ctgcgggacc	agcaggcact	gccggcactc	tggcgcgacc	tggatccgcg	55740
ggcccgccgc	accgcggcca	ccgcgcgacc	ctcccgggtc	acggtgcgcg	agcgccctcg	55800
ccacctcgac	gagacggggc	aggagcagct	gctcatcgac	ctcgtcgctg	gctacaccgc	55860
cggcctgctc	ggccaccccc	accccaccgc	cgtcgatccc	gaacgggggt	tcctggagct	55920
gggcttcgac	tccttgggtc	cggtcggcct	cgcaaccag	ctcgccgaga	tcctcggcct	55980
gcgcctgccc	tcgtccatcg	tcttcgacag	caagtcggcg	gtgaagctgg	cgcgttggct	56040
gcaccaggaa	ctcgccaacg	gccccagcc	gggcgccacc	ggccccggcg	ccgcggagcg	56100
ccgtcccgcg	gtgcgctccg	acgacacct	ggagggcctg	ttctacaacg	cggtgccggg	56160
cggcaagctc	gtcgaggcga	tgcggatgct	caaggccgtc	gccaacaccc	ggccgatggt	56220
cgcacacccc	gccgagctgg	aggagctctc	cgagccggtg	acgctcgccg	acggcccggg	56280
ccggccccgg	ctgatcttcg	tcagcgcccc	gggcgcacc	ggcgcgctcc	accagtacgc	56340
gcgcacgcgc	gcgcacttcc	gcggcagccg	ccatgtctcc	gcgctgcccc	tgatgggctt	56400
cgcccccggc	gagctcctcc	cggccaccag	cgaggccggc	gcccgtatcg	tcgccgagag	56460
cgtcctgatg	gccagcgagg	gcgaaccgtt	cgtcatggtc	ggccactcca	ccggcggtct	56520
gctggcctac	ctcgccggcg	gcgtcctgga	ggacacctgg	gacgtccggc	ccgaagcggt	56580
ggtcctcctc	gacaccgcgt	ccatccgcta	caaccccggc	gagggcaacg	acctggaccg	56640
caccacgagg	ttctacctgg	ccgacatcga	ctcgccctcg	gtgacgctca	acagcgcccc	56700
gatgtccgcc	atggcccaact	ggttcctggc	gatgaccgac	atccaggcg	ccgcaccgac	56760
cgcccccacc	ctcctcgtgc	gcgcgcggcg	ggccctcgac	ggcttccggc	tcgacacctc	56820
gtccgtcccc	gccgacgagg	tccgggacat	cgcgcggcgc	cacctctccc	tcgccaagga	56880
gcactcggca	ctgaccgcgc	aggccatcga	gggatggctc	gcggaaactgc	cggacccccg	56940
ggcctgatct	ccggccccgg	cggcccccca	caggccgggc	ccatcccccc	accggggcg	57000



cgccacccag	gtgcgccccg	gcgcaccccg	ccccgcgcag	gcagccggag	cgcccatccc	57060
cccgcattgg	accaccctcc	agaggagtc	ttccatgagc	acaccgaccg	caccgccttc	57120
cctgaaagcg	gaggtgccgc	ccgtcctgcg	cctgagcccg	ctgctgcgcg	aactccagtc	57180
ccgcgcccc	gtctgcaagg	tccgcacccc	cgccggcgac	gagggctggc	tggtagaccg	57240
gcacaccgaa	ctcaagcagc	tgctgcacga	cgaccggctg	gcccgcgccc	acgccgaccc	57300
ggccaacgcc	ccgcgttatg	tgcacaaccc	gttcctggac	ctgctcgteg	tcgacgactt	57360
cgacctggcc	cgcacgctgc	acgccgagat	gcgtccttg	ttcaccocgc	agttctcggc	57420
ccgcgcgctc	atggacctga	cgccgagggg	ggaagccctc	gccgaggggg	tactggccca	57480
cttcgctcgcc	cagggacccg	ccgcgcgacct	gcacaacgac	ttctcgctgc	cgttctccct	57540
gtcggtgctg	tgcgcgctca	tccggctccc	ggccgaggaa	caggggaagc	tgatcgccgc	57600
cctcaccaaa	ctgggcgaac	tcgacgaccc	ggcacgcgtc	caggaaggcc	aggacgagct	57660
gttcggcctg	ctgtccggcc	tggcacgcgc	caagcgcata	acaccgagg	acgacgtcat	57720
ctcccggctc	tgcctgaagg	tgcctccga	cgagcgcata	ggcccgatcg	cctccggtct	57780
gctcttcgcc	ggcctggaca	gcgtcgccag	ccacatcgac	ctgggcacgg	tgctgttcat	57840
ccagcaccgc	gaccagctcg	ccgcggccct	ggccgacgag	aagctgatgc	gcggcgccgt	57900
cgaggagatc	ctgcgggtccg	ccaaggccgg	cggttcggtg	ctcccgcggt	acgcgaccgc	57960
cgatgtaccg	atcggcgacg	tgaccatcag	ggccggcgac	ctggtgctgc	tggacttcac	58020
cctggtgaac	ttcgaccgca	cggctcttca	cgagccggag	ctcttcgaca	tccggcgcg	58080
ccccaacccg	cacctgacgt	tccggccacg	catgtggcac	tgcattggcg	cgccgctggc	58140
ccgggtcaat	ctgcgcaccg	cctacaccct	gctgttcacc	cgctgcccc	gcctgcggt	58200
ggtgcgcccc	gtcgaggaac	tgcgggtgct	gtcggggcag	ttgtcgcccg	gcctgacgga	58260
gtgccccgtc	acctggtgac	gtgatgtcgg	acccgccccg	cccggttccg	ggcaggcgga	58320
agtgaaggcc	cccgccccgc	tccggccggg	gcccacacgc	acgggggagc	gggtcctca	58380
gtccggccgc	acgctgatcg	ccccggacgg	gcagagtcgc	gcgcgctcgc	gtacgtcgcc	58440
gggatccgcg	gcgtccgcgc	ccccggccag	gacggtcacc	agaccgtcgt	cgctcgtggt	58500
gaacagggtc	ggtgcgggtca	ggacgcactg	gcccggcccc	acacagcggc	cgggggtccac	58560
ggtgatgcgc	acgatcggtt	ctccgagggg	cggttcgtcg	ccggggtcgg	gcccgcgcgc	58620
ggcggcccca	ccgggtcac	ttccagggtga	cgggcagttc	gtgcaggccg	aacagcacc	58680
cgctgtactt	caacggcagt	tctccaccg	gcacggcgag	cgtgagggag	gggatccggg	58740
cgaagagcgt	gcggtaggcg	atgtccatct	cctccgcac	caggttctgg	cccaggcaact	58800
ggtggacacc	gtagccgaac	gccacgtgcg	agcgcgccga	gcgtcgggg	tcgaactccg	58860
agggcgccgc	gaaggccgtg	gcgtcgtggt	tggcgggcgc	gatcagcggc	acgatgccct	58920
cgccggcccc	gatcagctga	ccgcgcatct	ccacgtcctc	gacggccacc	cggaatgcca	58980
ccagatcggc	caccgagtag	tagcgcagca	gttcctcgac	gatccggtcg	tcgccgatcc	59040
accgcgggtt	ggtgagcaac	tgcaccacgc	ccaggccgat	gttggttcgcc	gtcgtctcgt	59100
gcccggcgat	cagcagcagc	atcgccaccc	cggacagctc	ctgccgggac	agtgtgcccg	59160
cggcgatcag	ccggtgatc	aggtcgtcac	cgggtcgccg	ctgcttgatc	tggatgagtc	59220
ggcccagata	gcgcagcagc	gcctgggtcg	ccttgctcgc	ctcctcgteg	gtcgaactga	59280
gcccggaccag	caccgggtc	cggctcctga	agaagtcctg	gtcgaccttg	ggcacccca	59340
gcagactgga	gatcaccagg	gacggcaccc	gcagcgcgaa	ggactcgacg	aggtccgcgc	59400
aggtgccccg	ggccagcatc	gcgtcgatcc	gctcgtccac	gatctgctgg	atcaccggcc	59460
gcagttcccc	caccttgccg	acgggtgaact	cggggatcaa	ggctctgcgg	aaccggccgt	59520
gctccgggtg	gtccatcgcc	acgaaccacc	ccggcacctc	gtactgcgac	ggggcccccac	59580
cgggtccgacc	ggccttgggg	aaaccgggtt	tggagggggt	ggcgctgatc	cgggggtcgg	59640
tcagcaccgc	ccggacgtcc	tcgtgccggg	tcaccaacca	gaccggggccg	ctggggagtt	59700
cggagcggac	caggcccgcc	ccgcgcgggt	aggtcgcgta	ctccggcggg	gggaaggggc	59760
ggcccggcct	gcgcagcggg	aaggccaccg	ggcactgcgg	cgcggcggtc	cgcgcgtcgg	59820
cttcgggtgct	catgccactc	cgtagaacgc	gcggatccgg	ccggtgatga	actcctgctc	59880
ctgtgcgggt	agtccgggtg	gcgtcggcag	atagaagccg	tcggcgctga	gccggggccgc	59940
gttcaacgcg	ggccagtcgg	cggagaagta	accgggctga	cggctcatcg	gcttgaagaa	60000
gacccgggtc	tcgatgccct	cgccggcgag	gtacgcgcac	agctcgtccc	gccgctcggc	60060
ccgcaggctc	tacatccaca	gcacgtcgcg	cggcggcata	agcgtgatgc	cgggggatgtc	60120
gcgcagcgcc	tcgtcgtagc	gcttctcgat	gtcgcggcgc	agggcgagga	tgggtgtccag	60180
ctgctcgggt	tgcgcgagcg	ccaccgcggc	ctgcatgttg	gtcatccgga	agttgtaggc	60240
cagcttcttg	tgcaggaagc	tgtggtcctt	ggtgaacgcc	atcgcccga	gatgggccat	60300
ctgctcggcc	aggtgcgggt	cgtgggtcag	gcagacgcgc	ccctcgccgg	ccgagatgat	60360
cttggtggcg	aagagcgaga	aacaggcgat	gtcgcgccgc	ggccgcaccc	cgtgcgcctc	60420
ggcggagtcc	tccaccaccc	gcaggttgta	ctcgtacgcc	aggttcagca	cggcgtccat	60480
gtcgcactgc	cgggcgtaga	tgtgcaccgg	catgatcact	ttggtgcgcg	gggtgatctt	60540
ctcctcgatg	cgggacacgt	cgatgttcag	gtcgtcgccg	cagtcacga	acaccggcgt	60600

ggcaccggtg	taggtcaccg	cccaggcgga	cgcgatcatc	gtgaactccg	ggacgatcac	60660
ctcgtcaccg	gggcccagcg	ccagcgcgcg	cagcgccagc	gtcagcgccg	tgggtgccgga	60720
ggagcaggcg	acgccgaacg	gcacgtcggt	gtacgcggcg	aacgcctcct	cgaaccgcct	60780
gacgtacggc	ccctgcgaag	agatccagcc	gccgcccagc	gcctccgtca	catagtcgag	60840
ctcgcggccc	tggagccacg	gcatggacac	cggatacgta	aaggacatgg	gtttttgagtc	60900
ttcctcgggtc	agtcggttgc	cagggcgggg	aggccgagca	gcaggtccgc	cgccgcggcc	60960
cggccgcggg	cgtcccgcag	cagcccggcg	aagtgtcccg	cgcgctcggg	gaaggagggt	61020
tggtcgagca	cgcgggtgat	cttgtccagg	acgtcctcgg	tgtccacggg	ctccggccgg	61080
tccaggggtca	ggctcacccc	gaagtcctgg	ccccggatcg	cctgggtcgtc	gcagtcacc	61140
cacaacggcc	ggaccaccag	cggctttccg	aagtacaggc	cctcgtggta	gccgttgcca	61200
ccggcatggg	tgaagaacgc	cttcacgttc	ggatgggcca	gcacgtccag	ctgcgacggc	61260
accagaccct	cgatccgcag	gttgtccggc	agctcggcgg	ccggcggcag	caactcctgt	61320
tggccgcgcg	ggagtttcca	caacacctgg	tggccccggc	cgtccagtcg	ccgggcgacc	61380
tccaccagcg	acgccacctg	ctcacgggtc	agccgggtga	tctgtccgaa	gccccatgtac	61440
accacggact	tctgcgccga	cagccagtc	gacaggccgt	cgtcgtccgg	tgcctggggc	61500
agcggcgggc	ccatcgtgcc	caccagccgc	agcttcggat	gcacgggaa	cgggtagtc	61560
aactccctta	cggagtagca	caagacctgc	tccgcatggg	cgatccgcgc	catcatctgc	61620
cgcgcctggg	gcgcgatgcc	cagctcgggt	cggaccgggt	tgtcctcctc	gacgaccttg	61680
cggacgtccg	acgtcaggaa	catcccagag	gtccgcagcc	ggaacagctg	gttctcgatc	61740
cgctgagcca	gggacatcgc	ggccggcagc	cccagtgcg	gcaccgggaa	acccgacggg	61800
gtgtaggact	tggcgaacgg	gacgtgcgag	gtgaggacgt	tgtcggcac	gaacggcacc	61860
ccgagcacga	acggaatgcc	cttggtgatc	gccagctcgt	acccgaactg	gcacatgctc	61920
tcgatcacca	tcagcgccgg	ctcgacctcc	tcgacgatct	cctccaggcg	gcggtacttc	61980
gccatccgcg	actccggcgc	gaacgaatgc	cgaatcaccg	ccgcgtgcgc	cttgaaccgc	62040
gaccgctcgc	tacctccgc	atacgtcgcg	tgctcccatg	tgaccgccga	catctcgag	62100
acgggtgtcgc	cgagcgacgc	gaaccgaacc	gggctgccgt	ccaccacggc	cgccacctcg	62160
tcgcgcgctt	tctcgtcggg	ggcgaaccac	aggctccgcca	cgtcgcgccg	ggacaattcc	62220
ccggccagca	cgagcagcgg	attgagcagg	ccgctttcgg	cataactgac	gaacaggatc	62280
ggccgcccag	tcgcgcccag	ggacaacacc	cctcggaaatg	tggcgggccc	ccgggcccgc	62340
gcgccacgca	cccgcgccgg	ccggctcgccg	ggtagtgca	ttcgccgacg	ccgccacccg	62400
aggcgcgtgt	tgcgcgaagg	aagggtcacc	ggccggcacc	cggaaacgcgc	cgcgtggaaa	62460
acgggtcggg	tacttggtct	catgccacgg	accggggaat	cactagtctt	cggcgcgcga	62520
cggccctttc	cgggcccgtgt	ggccaatgcc	cgtccccggc	gcccgtcatt	ccttagggaa	62580
aagtacagcg	tttgcgaaacg	tacgatccgg	cacgcagagg	tgacctgagg	ccaacttttc	62640
cgcaggggtg	agcaaggcat	gacgatcgga	gccgacgagg	acccggtggg	ggtcgtcgga	62700
atggcctgcc	gttatccggg	tggggtcgcc	ggcccggagg	acctgtggga	actgggtccgc	62760
accggccgcg	acgcgaccac	cgccttcccc	gacgaccgcg	gctgggacct	ggccgcactg	62820
gccggcgacg	gacccggccg	cagcgcgacc	cgcgagggcg	gattcctcac	cggcgcgcc	62880
gacttcgacg	ccgccttctt	cggcatgtcg	ccccgcgagg	ccgtctccac	cgacccgcaa	62940
cagcgctcgc	tcctggagac	cgcctgggaa	gccctggagc	gcgcccgcac	cgacccgcac	63000
tcctgcgcgc	gcagccgcac	cggggtcttc	gtcggcgcca	gcggccagga	ctacgccgcc	63060
gtcaccacgc	cctcgcggga	cgacctggag	ggacacgccc	tcaccggcct	ggccccgggc	63120
gtcgctcccg	gtcgctcggc	gtacgtcctg	ggcctcgaag	gccccgcgct	caccgtcgac	63180
accacgtcct	cctcgtcgtg	ggtcgcgctg	cactgggcgg	tcgcgcgccct	gcgcgcgggg	63240
gagtgcagca	ccgcctggc	cggcggcgtc	acggtgatgt	ccaccccggc	cgccttcgtc	63300
ggccacaccc	gacaggggcg	cctcgcgccc	gacggccgct	gcaagccgtt	ctccgacgac	63360
gccgacggca	ccgcctgggc	ggaggggcgtc	ggcatcgctg	tcctggagca	cctgtccacc	63420
gcccgcgccg	ccggcaaccc	cgtcctcgcc	gtgctgcgcg	gctcggccgt	caaccaggac	63480
ggcgcctccg	acggcctcac	cgcacccagc	ggtccccgcc	aggaaacgcgt	catccgcgcc	63540
gcctcgcgcg	acgcccgcact	cgcccccgcc	gacatcgatc	tcgtcgaggc	gcacggcacc	63600
ggcaccgggc	tcggcgaccc	cgtcgaggcc	cgggcgctgc	tcgccgccta	cggccaggac	63660
cgggaccggg	accgaccgct	gcgcctcggg	tcctggaagt	ccaccctcgg	ccacgcacag	63720
gccgcgcgcc	gcacggcgcg	agtgatcaag	accgtcctga	ccctgcggca	cggcctgatg	63780
ccgcgcaccc	ggcacctggc	cacccccacc	cgcaagtcg	actggtccca	gggcgcgctg	63840
gccccctca	ccgaccacac	gccctggcca	ccggccgacc	gaccgcgccg	cgccggcgctc	63900
tcctccttcg	gcacagcgcg	caccaacgcc	catgtgatcc	tcgaagaggc	gccgcccgcc	63960
gacgtccccg	tcacccggcc	cggcaccttc	cgccccagca	ccgtccccctg	gccggtctcc	64020
gccgccacgc	ccgaagccct	cgacgcccac	ctcgccccgg	tcgcgcgcca	cctgcgcacc	64080
cactcggacc	tggaccggct	ggacgtcggc	tactccctgg	ccaccggccg	cgccgcgctc	64140
cggcaccggg	cggctcctcct	gccgcgccgc	gacggcaccg	ccgcggacgc	cgtcgagcac	64200



gcccgcggtg	cggcccacca	gcgcgcgacc	gcgcgtcctct	tctccggcca	gggcagccag	64260
cgcccgggca	tgggcgcgga	actgcgcgc	cgttcccg	tggtcgccga	cgcactggac	64320
gacgcgtgc	gcgccttga	ccggcacctg	gacggcccg	tgccgcaggt	gatgtggggc	64380
accgacgcc	cgtccttga	ccggaccggc	tggaccacgc	ccgccctgtt	cgccgtcgag	64440
gtcgccctcc	accgcctgg	cgcgtccctc	ggcgtcacc	ccgacttcgt	cgggggccac	64500
tccgtcggcg	agatcgccgc	cgcaccgtc	gcccggcgtc	tgagcctgga	ggacgcctgc	64560
cgcctgggtg	ccgcccgcgc	cacgtgatg	caggcgtcc	cgccggcg	cgcgatggcc	64620
gcgctggagg	ccaccgagga	cgaagtggcc	ccgtgctcg	gcgcacacct	cgcgtggcc	64680
gcccgtcaacg	gccccaccgc	ggctcgtcgc	gcccggagccg	aggacgccgt	gcccgaactg	64740
accgcccgt	tcgcccagccg	cgccggcg	accagccggc	tggccgtctc	gcacgccttc	64800
cactcgccgc	tgatggagcc	catgctcgac	gccttccggg	acgtcgtgag	ccgactgacc	64860
ttccaccagc	cgtcgatccc	gctggtctcc	aacctcaccg	gtgaactcgc	cggcagtgag	64920
atcaccagcg	ccgagtactg	ggtccggcac	gtccgcgaca	ccgtccgctt	cgccgacggc	64980
atcaccgcac	tggccaaggc	cgccggcgac	gtcctgatcg	aactcggccc	cgccggcggtg	65040
ctgtccgcga	tggcccgcga	cacctcggc	cccgcacagca	ccaccgacgt	cgtccccgcc	65100
ctgagcaagg	gacggcccga	ggagaccgcc	ttccggcg	ccctcggccg	cctgcacacc	65160
ctcgccgtcc	ccgtcgactg	gcccgccttc	tacgcggcca	ccggcgcccg	ccgcgtcgaa	65220
ctgcccacct	acgccttcca	gcacgtgcgc	cactggccca	ccccgccccg	cccgaacggc	65280
gcccggccccg	gcgccttcgg	ccaccccctg	ctcggtccg	ccgtcgaact	cgccgacggc	65340
ggcggcaccg	tctgctccgg	cgccctctcc	ctccgcaccc	acccctgggt	cgccgaccac	65400
accgtcgccg	ggcggttcgt	gctgcggg	accgcgtgc	tggaactcgc	cgtgcgcgcc	65460
ggcgacagg	cggtctgcga	cgtcctgcac	gaactccacc	tcaccacccc	gcccggccctg	65520
cccgacgacg	ccgccttcca	cgtccaggtg	cacgtcggcc	ccggcgacac	caccgggcgc	65580
cgcgcgtca	ccgtccacac	ccgccccgac	caccacccgg	ccggcgactg	gaccgatgc	65640
gccaccggca	ccctcggcag	caccccggcg	tccgcagccg	aagccgccac	ggcgccgacc	65700
ccggccgct	ggccgcgggc	cgacgcgcaa	ccctcgcacc	tcgccgacca	ctacgaacgg	65760
ctcgccgacc	gcggcttcga	ctacggcccg	accttccggg	gcctgcgggc	cgccggcgga	65820
cgcggcgccg	agatcttcgc	ggacgtggaa	tgcccgcccg	gcaccgcccga	cgacgcccc	65880
gaccacggac	tgacccccgc	cctgctcgac	gcggcccg	acgcgcgccat	ggcggtggac	65940
ggcaccgtgc	ccgtcgcctg	gcacggcgtc	cggctgcacg	ccgtcggcgc	caccgcgtg	66000
cgggctccga	tcgccccac	cacgaccggc	acgtgaccc	tcaccgcgg	cgacgtgcac	66060
ggcgcgccgg	tcgtcacctg	cgaggccctc	accgcccgc	cgctgaccga	cgaggaacgc	66120
gcccgcgccg	ggacgcgcg	gcaggcccg	ggcgagacgc	ccgcgcacgc	ccgcccggcc	66180
cgcccgcgg	cgcccgcgc	cgcccgggc	ggcgaacccc	tcgcccgcac	caccgggtcc	66240
caccccaccg	ccggccacct	cgcgcgctg	ccgcggccg	ccggggagcg	ccagctgctg	66300
gaactggtgc	gcacccaggc	cgcgcgcgtc	ctgggcccac	ccggccccga	ggcgcgtcggc	66360
accgcagcg	tcttcaaggga	gctgggttc	gaactcgttg	ccggcgctga	actcgcgcgac	66420
cggtcaccg	ccgcacccgg	actgcgcctg	ccggccaccc	tcgtcttcaa	cttccccacc	66480
cccgaacgtg	ctgcccaccg	cctcggggaa	ctcctcgccg	caaccgcccc	cctcgacccc	66540
ggggcgtag	gagaggaact	caccaggttc	gaggcgatcg	tgacgaacct	gcccagggac	66600
ggccccgaac	cgccggccgt	cgcggacccg	ttggagccca	tcgtctccgc	actccgcacg	66660
aactcgcctg	cagaggtgcc	ctcctcgag	gaggacatcg	acacggtgtc	ggtcgcagaga	66720
ctgctcgaca	tcacgatga	agagttcgaa	accacataga	gaaattgttg	cttctgctcg	66780
cgacccgatg	acgaggacgg	accgatgcag	gaacccagc	aaggccagcc	ggaccagcag	66840
gagaaaatcg	tcgactatct	ccggcgggtc	acttcagatc	ttcgccgtgc	ccgcgcgcgc	66900
attggcgaa	tggaaatcaa	ggacaacgag	cccatcgcca	tcgtcggaat	gggctgccga	66960
cttcccggcg	gcgtcaattc	gccggaatcc	ctgtgggacc	tggtgcgttc	cgccggcgac	67020
gccatttccg	gattccccgt	cgaccgcggc	tgggacctgg	aaacccctcac	cggaaacggc	67080
gacggcagca	gcgccaccca	cgaaggcgga	ttcctctacg	acgcgcggga	attcgacgcc	67140
gccttcttcg	gcattctgc	gcgcgaggcg	actgctatgg	accccccagca	gcgcctctc	67200
ctcgaagtgc	cctgggaggg	gctggagcgc	gcccggcatcg	ccccacagc	cctgcgcggc	67260
agccgggtccg	gcgtgttcgt	cggctcctac	cactggggcg	cgccctcggc	cgacgcgcgc	67320
accgaactgc	acggccacgc	cctgaccggc	accgcgcgca	gcgtgctgtc	cgccgcgcctg	67380
gcctacaccc	tcggcctcga	aggcccggcc	gtcaccgtcg	acaccgcctg	ctcctcctcc	67440
ctcgtcgccc	tgacactggc	ggcccagtc	ctgcgcgtcg	gcgaatcctc	gctcgcgcgtg	67500
atcgccggcg	tcacgatcct	caccgagccg	tcgtcttcg	tcgagttcag	cgcccagggc	67560
ggcctggcac	cggacggccg	ctgcaaggcg	ttctccgacg	ccgcgcacgg	caccggttgg	67620
gcccggggcg	tcggtgtcct	cgtcgcgcgag	cggctctccg	acgcgcagcg	caacggccat	67680
ccggtcctcg	ccgtgctgcg	cggctccgc	gtcaaccagg	acggcgccctc	caacggcctg	67740
accgccccca	acggccctc	ccaggaacgg	gtcatccagc	aggccctcgc	ccggaccggc	67800

ctgacccccg	ccgacatcga	cgcgcgtcgag	gcgcacggca	ccggcaccgg	gctcggcgac	67860
cccacgcagg	cccaggccct	gctcggccacc	tacggccagg	gacacacccc	cgaccagccg	67920
ctgtggctcg	gctccctgaa	gtccaacatc	gggcacaccc	aggcggccgc	cggcgtcgcc	67980
ggtgtcatca	agatggtcat	ggcgtcgccg	cacggccacc	tgcgcgcgac	cctgcacgcc	68040
gacgcgccct	cctcgcacgt	ggactgggtc	gccggatcgg	tacgcctgct	gaccgagggc	68100
cagcagtggc	cggagaccgg	acgtcccgcc	gcggcccgcc	tgtcctcggt	cggcatcagc	68160
ggcaccaacg	cgcacgccct	gctggaacag	gcacccacc	ccgcggacac	cgcggacgcc	68220
ggcgacgacg	ccgcgcccc	cgaaccggcc	ggcgcccgcc	ccgcgctgcc	ctggatcgtc	68280
tccggacact	ccccgcaggc	gctgcgcgac	caggcccgcc	ccctggccgc	cagggtcgag	68340
accgaccccc	cgctccgccc	ccaggacatc	gggcacaccc	tgcacaccgc	ccgcgccctg	68400
ctcgaacgac	gcgcgcgtcg	cgctcgcccc	gaccgcgcgg	aactcctcgc	ggctaccac	68460
gagttggccg	ccggccgggtc	cgcgaacgcc	gtcgtcgagg	gcctcgcgga	cgctcaggggt	68520
cggacgggtg	tcggtgtccc	cgggtcaggg	tgcagtgagg	tggggatggg	ggcccaactc	68580
ctcgatgagt	cggcgggtgt	cgcggagcgg	attgccgagt	gtgcggcgcc	actcgcgcgag	68640
ttcaccgact	ggtcgctggg	cgatgtgctg	cgggggtgtg	tgggtgcgcc	gtcgttggag	68700
cgggtcgatg	tgggtgcagcc	ggcgtcggtc	gcgggtgatg	tgtcgttggc	tgcgttgtgg	68760
ggttcccgtg	gtgtgttgcc	ggatgcgggtg	gtggggcatt	cgcagggtga	gatcgtgcc	68820
gcgggtggtg	cgggtgcgct	gtcgttgccg	gacggggcgc	gggtggtggc	gctgcggagt	68880
caggccattg	gtcgtgcgtt	ggcggggcgg	ggcgggatga	tgtccgtcgc	gctgtcgggtg	68940
gacgtgctcg	aaccgcgggt	ggtcgagttc	gaggggcggg	tgtcgggtggc	cgccgtcaac	69000
ggcccgcgct	cgtcgtgggt	cgcgggcgag	cccagggcgc	tggacgcgct	gcacgcccg	69060
ctgaccgcgg	acgacatccg	ggcccgcggg	atcgcgggtg	actacgcctc	gcactcgcac	69120
caggtcgagg	acctgcacga	ggaactgctg	gaggtgctgg	cggagctggc	gccgcgcacg	69180
tccgaggtgc	cgttcttctc	gaccgtgacc	ggcgactggc	tggacaccgc	gcggatggac	69240
gccggctact	ggttccgcaa	cctgcgcgga	cgggtgcggg	tgcgcggacg	ggtggcggac	69300
ctgctggcgg	cggagtaccg	cgcgttcgtc	gaggtcagct	cgcacccggg	gctgacgatg	69360
gcgggtcttg	acctgatcga	ggaggccggg	gtcacggccg	tgcgcaccgg	cacctgcgc	69420
cgtgaccagg	gtggcgcggg	ccgcttcctg	ctgtcggccg	ccgaggtctt	cgtgcgcggg	69480
gtggacgtgg	actgggcggg	ggcgttcgag	gggaccgggtg	cgccccgggt	cgacctgccc	69540
acctacgcct	tccagcgcga	gcgggtactg	aacacccgca	ccgcgcgcga	ccgcaccccg	69600
gccgacgccc	cgatggacgc	cgaattctgg	gccgcgcgtg	aacaggcgga	cgtctccgcg	69660
ctgaccgcgg	cgctcggcac	cgacgaggac	tccgtcgcgg	ccatcctgcc	cggcctcacc	69720
tccctggcgg	gggcccgcgtc	ccagcgcacc	accctcgact	cctggcgcta	ccgcgtcacc	69780
tggacgcccc	tgcgccaggg	gccccgcgcc	accctgaccg	gcacctgggt	gctgggtcacc	69840
accgacggca	tgcacgacac	cgatgtggca	ggggcggttg	agagctacgg	cgcgcagggtg	69900
cgcgcggctg	tccctggacga	ggagtgcacc	gaccgcgcgg	tccctgcggga	gcggctggcc	69960
ggcgcgaggg	acgtgaccgg	catcgtctcc	gtcctcgcgg	ccgcgcaggga	cgacgccgca	70020
cgccaccccc	gcctcaccgg	gggactcgcg	ctcaccgtct	ccctcgtcca	ggccctgggc	70080
gacgcgcagg	cgaccgcgcc	gctgtggttc	ctgaccgcgg	gcgccttcgc	caccggcccg	70140
tccgaccccc	tcaccgggcc	cctgcagagc	cagatcgcgg	gcgtcggctg	gaccaccgcg	70200
ctggagcacc	cgcagcgtcg	gggcggcacc	gtggacactgc	ccgacaccct	cgacgcccg	70260
gccgcccagc	ggctcgcggc	cgcgctgtcc	ggcgccctcg	gcgcgcaggga	ccagctcgcc	70320
gtccgcgcgg	cgggggtact	ggcccgcggc	atcgtgcgtg	ccggacaccg	cgcgcggacga	70380
ccggcacgga	cctgggcggc	gcgcggcacc	accctgatca	ccggcggttc	cggcacccctc	70440
gccccgcagc	tgcgccgctg	gctggccgaa	cgcggcgccg	agcacgtggg	gctgggtcagc	70500
cggcgcggtg	ccgacgcccc	cggagcgccc	gaactcatcg	cggaggcagc	cgagtcgggc	70560
accgaggtga	ccgtcgcggc	ctgcgacatc	accgaccgcg	acgcgggtcg	cgcgctgctg	70620
gccgacctca	cggccgacgg	cgcacccctg	cgcaccgtca	tccacgcgcg	cgcgcgccatc	70680
gagctgtccg	cgctcgcgga	caccaccgtg	gcggagttcg	ccgacgtcgt	gcacgccaag	70740
gtcaccggcg	cacggatcct	cgacgaactg	ctcgacgacg	cggaaactgga	cgacttcgtc	70800
ctgtactcct	ccaccgcggg	catgtggggc	agcggcggtg	acgcgcgcta	cgtcgcgggc	70860
aacgcctatc	tgtccgcgct	cgcgcgagcag	cgcgcgcggc	gcggactgcg	caccacctcc	70920
atccactggg	gcaagtggcc	cgacgaccgg	gcacgcgagc	tggccgaccc	gcaccggatc	70980
cgcgcgcagg	gtctggagta	cctcgacccc	gagctggcgc	tcaccgcgct	ccagcacgtc	71040
ctggacgacg	acgagaccgt	catcggcctc	atggacatcg	actgggacac	ctaccacgac	71100
gtgttcaccg	cgggcggggc	cgcgcacctc	ttcgaccaga	tccccgaggt	gcggcgccgc	71160
ctcgaccagg	catccgtccc	ggaccccgcg	ggcccgccgg	ccgacggcct	ggccgcggcg	71220
ctgcacggcc	tgcgcggcgc	cgaacaggac	cggctgctgc	tcaccctggt	ccgcaccgag	71280
gccgcgcgg	tccctcggcca	cgcctcgggc	gagtccttcc	ccgagcgccg	cgccttcgct	71340
gacctcgggt	tgcactcggg	caccgcgggtg	gacctgcgca	accggctcgt	ggccggcacc	71400

ggactgcggc	tgccctcgac	gatggtcttc	gaccacccca	actgcgcggc	gctcgccgcg	71460
ttcctgaaga	cgacggcgct	cggcgtcccc	ggcgccgcac	cgacgcagca	cgccgctacc	71520
ggcaccgccg	ccgacgacga	cccgatcgcc	gtgatcgcca	tgagctgccg	ctaccccggc	71580
ggcgccgccca	ccccgagga	actgctgcgg	ctcgccctcg	acggcgccga	cgatcatctcg	71640
gagttccccg	cggaccgcgg	ctgggacgcc	cggggcctgt	acgaccggga	ccccgaccgc	71700
ccccggccaca	cctactccgt	ccagggcggc	ttcctccacg	aggccgcccg	cttcgatccc	71760
ggcttcttcg	ggatctcccc	gcgcgaggcg	gtcgccatgg	acccgcagca	gcggctcctg	71820
ctggagacct	cctgggaggg	gttcgaacgc	gccggtatcg	accccgcgtc	actgcgcggc	71880
agcgccgccc	gcaccttctt	cggcgccagc	taccaggact	actcctccac	cgtgcagaac	71940
ggcacggggg	agtcgaggg	gcacatgggt	accggcaccg	cggccagtgt	cctgtccggc	72000
cgggtctcct	acctgctcgg	cctggagggc	cccgcggtca	ccgtggacac	cgctgtctcc	72060
tctcactgg	tcgccctgca	cctggcctgc	cagtcctctg	gcgacggcga	gagctccctc	72120
gcgctggccg	gcggtgcggc	cgtgatggcc	acccgcacg	cgttcgctcg	cttcagccgg	72180
cagcgtgccc	tggccaagga	cggccgctgc	aagccgttct	ccgacaccgc	cgacggcatg	72240
acgctcgccg	agggcgctcg	cgtcgctcct	ctggagcgcc	tgtcccacgc	ccgcgccaac	72300
gggcaccggg	tactggccgt	gatccgcggt	tccgcgctca	accaggacgg	cgctccaac	72360
ggcctgaccg	cgcccaacgg	cccgtcccag	cagcgcgctca	tccgccaggc	gctcgccaac	72420
gccggcctga	ccggcgccga	cgtcgacgcg	gtcgaggcgc	acggcaccgg	caccaagctg	72480
ggcgacccca	tcgaggccca	ggccctgctc	gccacctacg	gccaggaccg	cgacgccgaa	72540
cggccgctgc	tgtctgggctc	ggtgaagtcc	aacatcggcc	acaccaggc	cgcgcccggc	72600
gtcgccggcg	tcataaagat	ggtgctggcc	atggacgcgg	gcgaactgcc	cggcaccctg	72660
cacctcgacg	cgccctccag	ccacgtcgac	tggaccgcgg	gcgccgtgga	actgctgcgc	72720
gggcgcaccc	cgtggcccca	gagcgggcgc	ccccgcgggg	ccggtgtctc	ctcgttcggc	72780
atcagcgga	ccaacgcccc	cctgatcctc	gaacaggccc	cggccaccga	gccgcagcc	72840
gaccccgacc	gcctccggga	caccgccacc	gacacgtcgc	tccctggcc	gctcgccgc	72900
aagtcccccg	cgccctgcg	cgcccaggcc	gcccggctcc	tcgccaccgt	cgagcacgac	72960
cccgaacctc	cgcccgcccc	cgtggggccac	gcccctggcca	ccaccgcgc	cgccctcgaa	73020
caccgcgccc	tcgtcgctcg	cgagcgccgc	gaggacttcc	tcgcgcggct	ggccgccttg	73080
tccaccggcg	cctcgacggc	cgccctgggtc	agcggcatcg	ccggccccga	ccccgaggga	73140
gcggtcttcg	tcttcccccg	ccagggatcc	cagtggtggg	gaatgggccc	cgaactcctc	73200
gccacgtccg	aggtgttccg	caccgcgatc	gatgactcg	cgacggccct	cgcccgctac	73260
gtcgactgg	cgctgcacga	cgctcctggc	ggcgagggcg	accccgccct	gctggagcgg	73320
gtggacgtgg	tccagcccgc	gctgttcgcc	atgatggctg	ggctgtccgc	gctctggcgc	73380
tcccacggcg	tcgtcccggc	ggccgtgggtc	ggccactcgc	agggcgagat	cgccgcggcc	73440
tgctcgccg	gagccctcag	cctggccgac	gccgcccgcg	tgggtggcgt	gcgcagccag	73500
gactgcccgc	aactgtccgg	acgcggcggc	atgatgtcgg	tctccgcccc	cgtagagcgg	73560
gtcaccgcac	tcctcgcccc	gtggcaggag	gcgctgtccg	tcgccgcgggt	caacggcccc	73620
tcgtccgtgg	tcgtctccgg	cgacaccgac	gcgctcgacg	ccctgcacac	cgctgccag	73680
gaacagggcg	tcggggcccc	caaggtgtcc	gtggactacg	cctcgcacgg	gcggcacgtc	73740
gaggccgtcc	gcgacgaact	cgcccgcgtc	ctcgcgccgg	tcgaccgcgg	cgcccccgag	73800
gtgccgttct	actcgacgg	caccggcgac	cgctggacg	acggcgccct	cgacggcgcc	73860
tactggtaca	ccaacctccg	ccagaccgtc	cgcatggagg	aggccaccgg	cgccctcctc	73920
gccgccggac	accgcgtctt	catcgaggtc	agcccgacc	cggtgctcgc	cgccccgatc	73980
caggagacgc	aggaggccgt	agcggaggcc	accggcggg	ccgcgggtgt	cctcgctcgc	74040
ctccgcgcg	acgaggcg	cccgcggcgc	ttcctgacgt	cgctcgccga	ggcccacacc	74100
cacggcgccc	cgtcgactg	gaccaccacc	ttcgcccggt	ccgcctacca	gccggtggac	74160
ctgcgcacct	acccttcca	acgacaggac	ttctggcccc	aggcccgggc	cgccaccccg	74220
gccgcggcg	ccgacgcgtc	cgacgcgcg	ttctggcaac	tggctcgagaa	ccaggacctc	74280
gccgcgctcg	ccgacgcgtc	cggcgtcccc	gccgacgacg	agcacaccgc	gctcggcacc	74340
gtgctgcgg	ccctgtccgc	ctggcgcgcc	aaggcccagg	cccgaccccg	gatcgacgaa	74400
ctccgctacc	acgtccagt	gaccgggtc	gccgagcccg	cggcggcccc	caccaccggc	74460
cggctgctgg	tcgccgtccc	gccggaccac	gccgacgccc	cctgggtcgc	cgcgcgctc	74520
gacgccttg	gcaccgacac	cgtccgcttc	gaggccaagg	gcaccgaccg	cgcgggatgg	74580
gccgcacaga	tcgcccgaact	cgtcgaggac	ggcgaggagt	tcaccggcgt	ggtgtcgtg	74640
ctggccgccc	ccgaggatct	ccaccgggac	ttcggtcctg	taccgctggg	gctggggcag	74700
accctcgtcc	tcgtccaggc	cctcggcgac	gccggcctga	ccgcgcccct	gtggtgctg	74760
acccgcggcg	ccgtcgccac	cggccgcgac	gacgcctcgc	acagcccga	ccaggcgcc	74820
ctgtggggcc	tcggccgggt	cgtggccctg	gaacaccccc	accgctgggg	cgccctgatc	74880
gacctgccc	ccacctcga	cgcccgcgcc	gcggcccgcc	tcaccggcct	gctcgccgac	74940
cccgcgggtg	aggaccaact	cgccgtccgc	gccaccggcg	tgctcgcccc	ccgcatgggt	75000

cacgccgcgc	cgccgcgcgc	ccgcaccggg	cgccgcctggc	ggggccgcgc	cacctgcctg	75060
atcaccggcy	gcaccggcgc	catcgggcgc	cgggctgcgc	gctggatggc	cgagcacggc	75120
gcgcgccacc	tggctctgac	cagccggcgc	ggcccgagc	caccggcgc	cgccgcactc	75180
cgggccgaac	tggaggccct	gggcgcgcgc	gtcaccctcg	ccgcctgcga	cgccgcgcac	75240
cgcgacgccc	tggccgcctt	gctggccgac	ctccccgcgc	accagccgct	cacctccgtc	75300
ttccactccg	ccggcgtggc	cgacggggac	gcccgggcag	ccgacctgac	cctagatcag	75360
ctcgacgcgc	tgtgcgcgcc	caaactgacc	gcccgcacc	acctgcacga	gctgaccgcc	75420
cccctcgacc	tcgacgcgtt	cgtgctcttc	tcctccggcg	ccgcggtctg	gggcagcggc	75480
ggccagcccc	gctacgcgcg	cgccaacgcc	tacctcgacg	ccctcgccgc	ccaccgcagg	75540
tcctctgacc	tggccggcgc	gtccgtcgcc	tggggcagct	ggggcgaggt	cggcatggcc	75600
accgtccccg	aggtccacga	gcgactgcac	cgccaagggg	tcgcgcgcct	ggaaccggac	75660
cacgcgatcg	gcgcgctcca	gcagatgctg	gaggacgacg	acaccacctt	cgccgtgacc	75720
ctcatggact	gggaggcggt	cgacccgagc	ttcaccgcga	ccgacccag	cgccctgttc	75780
agcacggtgc	ccgaagccgt	ccgcgcgggt	accggcgacc	cgggcaccac	ggccggcgac	75840
gacgtggact	ccgcgacccc	gccgtccgc	cgccacctgg	aggagctgtc	cgccgcgcag	75900
cgcgcccggg	ccctggctga	ggcggtccgc	gccgaggcgt	ccgcgacctt	cgccacacgac	75960
acccccgacg	ccatccccgc	cgcccggtgc	ttccgcgacg	tcggcttcga	ctcggtcacc	76020
gcgctcgaac	tgcgcaaccg	gctgcgccac	gccctcgccc	tggcgctgcc	ggccgcgctc	76080
gtcttcgacc	acccaccccc	cacggccctc	gccggccacc	tcggcgcgct	gctcttcggc	76140
accgcccccg	aggacgccgc	caccggccgc	cccgcgacc	ccgacgccc	catccgcgag	76200
gcgctcgcca	ccgtcccat	cggacggctg	cgcaaggcgg	gcctcctcga	catggtgctg	76260
aaactcgccg	acggagacgc	gaccgacgcc	cccgcccccg	aggccgacgc	cccctcgga	76320
tcctctgacg	acatggacgc	cgaagccctg	ctgcggttgg	ccaccgagaa	ctcggcgaac	76380
tgaagagag	ctggagccca	ccatgagcac	gaaccccgac	aagtacgtcg	aggcactccg	76440
gtcgctccctg	aaggagatcg	agcggtcgcg	ccggcagaac	gaacagctgg	tggcgccggc	76500
ggctcgagccc	gtcgcggtcg	tcggcatcgg	ctgcgccttc	cccggcgcg	tcacctcccc	76560
cgaggacctg	tgggagctgg	tcgcccagg	gcgcgacgtc	atcgggccgt	tcggcgagga	76620
ccgcggctgg	gacctggaga	agttggccgg	cggcggcgag	ggcgccagcc	tcgcccaggt	76680
cggcggcttc	gtcgaggacg	ccgcggcctt	cgaccccggc	ttcttcggca	tctccccgcg	76740
cgaagcggtc	gccatggacc	cgacgcgagc	catcctgctg	gagatcacct	gggaagccct	76800
ggaacgcgcc	ggcatcgacc	cgtccacctt	gcgcggtacc	cccaccggcg	tcttcgtcgg	76860
caccaccggc	caggactacg	gcgaggtcat	caaggcgtcc	gccgaggacg	tcgaggtcta	76920
ctcgaccacc	ggccacgcgc	ccagcgtcat	ctccggccgg	ctctcctaca	ccctcggcgc	76980
cgagggcccc	gccgtcaccg	tcgacaccgg	ctgctcctcg	tccttggctg	ccctgcactg	77040
ggccgtccag	gcgctgcgcg	gcggcgagtg	ctccatggcc	ctggccggcg	gcgcgtccat	77100
catggccacc	ccgggcccgt	tcgtcgccct	caccgcgcag	agcgccctgg	ccgccgacgg	77160
ccgctgcaag	cccttctccg	accgggcccga	cggcaccggc	tggggcgagg	gcgccggcat	77220
gctggtcctg	atgcggctct	ccgacgcccc	gcgcgagggc	cgcccggctc	tcgcccgtgt	77280
gcgcggctcc	gccatcaacc	aggacggcgc	ctccaacggc	ctgaccgctc	ccaacggccc	77340
ctcccagcag	cgcgtcatcc	gcgcgcgcgt	ggacagcgcc	cacctcaccg	ccgccgacat	77400
cgacgcgctc	gaggccacag	gcaccggcac	cacctcggc	gaccgatcg	aggcccaggc	77460
gctcctggcg	acctacggac	aggaccggcc	gcggcccttg	tggctcggct	cgggtgaagtc	77520
caacatcggc	cacaccagg	ccgcctccgg	tgcgcgcggc	gtgatcaaaa	tgatcatggc	77580
gttgacgcgc	ggcgtgctgc	cgcgacgctt	gcacgccacc	gaacccacca	cggacgtcga	77640
ctggaccgcc	ggctccgtcg	acctcctcga	cgagacggtc	gcctggcccc	agaccggacg	77700
cgcccgccgc	ggcgcgctct	cctccttcgg	catcagcggc	accacgccc	acgtcatcct	77760
cgaacaggcc	cccaccgccc	ccgaagagcc	caccaccgaa	cccaccgtcc	gccccgcgct	77820
cgtcccgtgg	gcgctctccg	cccgcaccgc	cgccgccttc	gacgcccagc	gcgcccgcct	77880
caccggccac	ctcgccgaca	ccccgcagcc	cgaccccttc	gacgtcggct	acgcgtcgc	77940
cgacggacgc	gccaccttcg	aacaccgcgc	cgtcctgctc	cccgcggcca	ccgaactcgc	78000
ccacggaacc	ggcggcgaag	gcccctgcgc	cgtcctcttc	tcgggccagg	gctcccagcg	78060
cccgggcatg	ggacgcgaac	tcacgcgccg	cttcccggtg	ttcgccgcgg	ccttcgacga	78120
gatcacagcg	ctcctcgaca	cccacctcga	ccgcccgcgt	cgcgaggctg	tctggggcac	78180
cgacgccgac	ctgctgaacg	acaccggctg	ggcccaaccc	gccctgttcg	ccgtcgaggt	78240
cgccctctac	cgcttggctg	cgtccctcgg	cgtcaccccc	gacttcgtcg	gcggccactc	78300
catcggcgag	ctcgccgcgc	cgacgctcgc	cggggtcctc	tcctcgaag	acgcctgcac	78360
cctcgtcgcc	ggccgcgccc	gcctcatgca	ggccctgcgc	cgcgccggcg	cgatgctcgc	78420
gatccgcgcc	accgaggacg	aggtcacccc	ccacctcacc	gacgacgtct	cgatcgccgc	78480
cgtcaacggg	cccacctccg	tcgtcgtcgc	cggcaccgag	gaagccgtcg	ccgcgatcgg	78540
ggcgcgcttc	accgcccagg	accgcaagac	caccggcgtg	cgggtcagcc	acgccttcca	78600

ctcgccgctc	atggacccga	tgttgccgga	attccgcgcc	gtcgccgcgg	gcctgacct	78660
ccacgagccg	cgcaccccg	tctctccaa	cctcacccgc	accgtcgccg	ccgtcgccga	78720
cctgtgctcc	gcccactact	gggtccgcca	cgtccgcgag	gcgggtccgt	tcgcccagcg	78780
cgtcacccgc	ctcacccgac	gcggcgtgac	cacgtctgtc	gaactcggcc	cggacggcgt	78840
gctgtccgcc	atggcccagg	aatccctgcc	ggacggcgcc	gccgcccgtg	cgtctgtcgg	78900
caaggaccgc	cccaggagc	tctccgcgt	caccggcctg	gccgcgcgcc	acgtccgcgg	78960
cgtcacggtc	cgtctggccg	gcctcttcga	cggcacccgc	gcgcgcgcgg	ccgacctgcc	79020
cacctacccc	ttccagcacc	agcggttctg	gccgaccgcg	gccgcgcgcc	cccaggacgt	79080
caccgcgcgg	ggactgggcg	ccgcgcgacca	cccgtctgtc	ggcgccaccg	tcgaactcgc	79140
cgcgggggccc	ggctacttgt	tcaccagccg	gctctccgtc	cggacccacc	cctggctcgc	79200
cgaccacggg	gtccaggggc	gggccctgct	gcccggcacc	gccttcgtcg	aactggcgt	79260
ccgcgcgcgg	gacgaggccg	gctgcgaccg	cgtcgaggaa	ctgacctgg	ccgccccct	79320
ggtgctgccc	gagcgcggcg	gcgtccaact	ccaggctcgc	gtcggcgccc	ccgacgcgcg	79380
cggccgcgcg	accctcggca	tcttctcccg	cgtcgaggac	ggcttcgacc	tgccctggtc	79440
gcaacacgcc	accggcgtcc	tgaccgcgcg	cgcggcgccc	cccgaaccca	ccttcgacgc	79500
caccgtctgg	ccccccagcg	gcgcgcgaacc	cgtcgacctc	accggcgcg	acgagcgct	79560
ggccgcactc	ggcttccagt	acggccccgc	cttccagggc	ctgcgcgcgg	cctggcgccg	79620
cgcacccgag	gtctacgccg	aagtggccct	gcccgacggc	gcggacaccg	accccgccgc	79680
cttcggactg	cacccggccc	tgttgacgc	cgcacaacac	gccgcgcgct	acgccgacct	79740
cggcgccatc	agccgcggcg	gcctgccgtt	cgcctgggaa	ggcgtctcgc	tcgcccgcgc	79800
cggcgccacc	accgtccgcg	cccggatcgc	cccggcgccg	gaggacaccg	tcaccatcgc	79860
cgtctacgac	gcgcgcggcg	gcaccgtgct	gtccgtcgac	tccttggctc	cccgcgaggt	79920
ccccgcgcgac	gcacccggcg	ccgcgcgcac	cgtccaccgc	gactccctct	tcacgtcga	79980
gtggaccccc	ctccaggggc	gcccggggccc	gcacccggcc	accgtcgccg	tcctcgccc	80040
cgcacccggac	gcccctcgccg	acaccctccg	cgcacccggc	atccggacca	ccgccccccg	80100
cgcacctggc	gcccctcgccg	acgcgcgaagg	gcccgctccc	gacctggctg	tcaccacct	80160
caccaccacc	ccgggcgcgc	ccgtccccga	cgcgcgcgac	gccaccaccg	ccgcgcgtct	80220
cgcctcgcgc	caacagtggc	tcgcccagca	ccgttcgcgc	gacgcccgc	tggtcctcgt	80280
cacccgcggc	gccaccgacg	gcaccgaccc	cgcgcgcgcg	gccgcgcggc	gcctgatccg	80340
caccgcccgc	accgagaacc	ccggccgttt	cgcctcctc	gacctcgccc	ccgacaccgg	80400
ccggcccgcac	cccagacccc	tgggccaccgc	cctggccgc	agccacgacg	agcccgacct	80460
cgcctcgcgc	ggcaccgacg	tgacgcgcgc	ccgcctggcc	cgtgtcccgc	tcgccaccga	80520
acccaccacc	tggaaccccg	acggcaccgt	cctgatcacc	ggcggcaccg	gcggcctggg	80580
cgcggtcctc	gcccgcacc	tggtcgccac	ccacggcgtc	cgcacactgc	tgctcgccag	80640
ccgcgcgcgg	ccggccgcgc	acggcgccga	cgcacctgacg	gccgaactca	ccgggctcgg	80700
cgcacccgtc	cacatcgccg	cctgcgacgt	cgcgcacccg	gccgcctcgc	ccgacctgct	80760
cggcaccgtc	ccggccgggc	acccgctcac	cgtcgtcgtc	cacaccgcgc	gcgtcgtcga	80820
cgcgcgcgc	ctcggtctcc	tcaccccgca	gcgcctggac	accgtcctgc	ggcccaaggc	80880
cgtctcgtcc	gtcgcgcgca	ccctcggcag	cccggacacc	gacctggacg	ccttcgtcct	80940
cgccttcttg	gacgcccctg	ccgcgcggcg	cgcgcgcaac	ggcctgcccg	ccacctcct	81000
cgcctggggc	ccgtggaccc	agagcgtcgg	catgacaagg	agcctgtccg	acctcgacgt	81120
cgcgcgcgac	gcccgcctcg	gcatgcccc	gctgacctg	gaacagggca	ccgcccctct	81180
cgcgcgcggc	ctggcgcccg	ggcccgcgcg	cctcgccccg	gtccgcctcg	acctgcccgt	81240
cctgcgcacc	cagggcgaca	tcgccccgct	gctgcgcggc	ctgatccgca	cccccgtcg	81300
gcgcaccgcc	gcccaggctc	cgcagaccgc	cgcaggcctc	gcccagcggc	tcgcccgcct	81360
cgcgcgcgc	gcccgcgggg	aagccctcct	ggaactcgtc	cgcacccaga	tcgcccagg	81420
cctcgccac	gcggacgcca	ccgagggtga	gaccggccgc	cagttccagg	acctcggtt	81480
cgcctcctc	accgcgcgtg	aactccgcaa	cgcctgaac	accgccaccg	gcctgcggct	81540
gcccgcacc	atggtgttcg	actaccgcac	accacacgcc	ctcgccgacc	acctgcgcga	81600
cgaactcctg	ggcaccgagg	ccgagtcgac	caccgcgcgc	cccgcgccga	cccgtaaccg	81660
cggcaccgac	gacccgatcg	tcacgtcgcg	catggcctgc	cgtacccccg	gcggcatcgc	81720
ctcacccgag	gacctctggc	gcctggctcag	ccaggcgccc	gacgccactg	gcccgttccc	81780
caccaaccgc	ggctggggacc	tggaacaacct	ctacgacccc	gaccccgacc	gcccggggcg	81840
caccacgcgc	cgcgcgcggc	gcttcctgca	cgcgcgcggc	tccttcgacg	ccgacttctt	81900
cgggatgagc	ccgcgcgagg	cgatggccac	cgcctcccag	cagcgccctg	tgctcgaact	81960
ctcctgggaa	gcccgcgaac	gcgcgcggcat	cgcacccgcc	tcactgcgcg	actccggcac	82020
cggcgtcttc	gcccggcgtc	tgtacaacga	ctacggcacc	accctgaccg	gcgacgagta	82080
cgcggcgttc	cgcggcaacg	gcgcgcgcgc	gagcgtcgc	tcgggcgcgc	tctcctacac	82140
cctcggcctg	gaaggccccg	ccgtcacggt	ggacaccgcc	tgctcttctc	ctctggctgc	82200

cctgcactgg	gcgggcgagg	cgttgcgggc	gggggagtg	tcgttgccgt	tggccgggtg	82260
tgtgacgggt	atgtcgacgc	cgagcacgtt	cgtggagttc	tcgcggcagc	gggggtctgg	82320
gcctgatggt	cgttcgaagg	cgttcgccga	ggccgcggac	ggcgtggcct	ggtccgaggg	82380
cgtcggcatg	ctggtcctgg	agcggcagtc	ggacgcgggt	cgcaacggtc	acgagatcct	82440
ggccgtgggt	cgcggctcgg	cggtcaacca	ggacgggtgc	tccaacggtc	tgaccgcgcc	82500
caacggcccc	tcccagcagc	gggtgatccg	tcaggcgttg	gccagtggcg	gcctgtccac	82560
ggccgacgtg	gacgccgttg	aggcgcacgg	cacgggtacg	acgctcgggt	accgatcgga	82620
ggcccaggcg	ctcctggcca	cctacggtcg	cgaccgcgac	cccagagaacc	cgctgctgct	82680
cggctcgatc	aagtccaaca	tcggtcacac	ccaggcagcg	gccggtgtcg	ccggtgtcat	82740
caagatgggt	atggcgatgc	ggcacggcgt	gctgccgcag	accctgcatg	tcgacgcgcc	82800
gtcctcgac	gtcgattgga	gcgtcggcgc	cgtcgaactg	ctcaccgagc	agaccgcctg	82860
gccggagacc	ggccggggcc	gtcgcgcggg	tgtctcctcc	ttcggcatca	gcggcaccaa	82920
cgcacacgtc	gtcatcgagc	agtccccgac	cgcgcgtccc	gccacgcccg	cgtccgcgga	82980
ccggtccgtc	gaggaaccgc	cggccgtccc	ctgggcccctg	tccggcaaga	cccccgacgc	83040
cctccgcgac	caggccgccc	gcctcctcgc	ccacgtcgag	gccacccccg	cactgcgccc	83100
cgtcgacatc	agctactccc	tgatcgccac	ccgcaccgcc	ttcgaccacc	gcgcgcgtcg	83160
cctcggcacc	gaccgcgcgg	aggccctgcg	cgcctcacc	gccctcgccg	ccggcgagac	83220
cgacccggcc	gccctcaccg	gcaccgtccg	caccggccgc	accgccttcc	tcttctccgg	83280
ccagggtccc	caacggctcg	gcattggcgg	cgtcctctac	gagcgggtcc	ccgccttcgc	83340
cgaagccctc	gacaccgtcc	tcaccgcctc	cgacgcggaa	ctcggccacc	ccctccgcga	83400
catcatcttg	ggcgaggacg	ctcaactcgt	cgaccggacc	ggctacaccc	aaccgcctc	83460
gttcgccatc	gaggtggcac	tcttcgcctc	ctggaagcc	tggggcatca	caccgactt	83520
cgtggccggc	cactccatcg	gcgagatcgc	cgcgcacac	gtcgcggcg	tgctctccct	83580
cggcgacgcc	tgccgcctcg	tcgtggcccg	cgcgcgtgct	atgcagtcgc	tgcccgaagg	83640
cggcgcgatg	atcgccgtcc	aggccaccga	ggacgaggtc	ctgcccctcc	tcaccgacga	83700
cgtctcgatc	gcccgcgtca	acagcccagc	ctcgcgtcgc	gtctccggct	acgagaacgc	83760
caccctcgcc	gtcgcgcggc	acttcgcgga	ccagggcgcg	cgcaccacgc	ggctgcgcgt	83820
cagccacgcc	ttccactcgc	cgtgatggc	gccgatgtc	gacgacttcc	gcgcgcgtcg	83880
cagagacctc	accttcaccg	ccccacgac	ccccgcgtc	tccaacctga	ccggcgaact	83940
ggccccggcc	gaggcgctct	gtcggccga	ctactgggtc	cggcacgtcc	gcgaggcggt	84000
ccgcttcgcc	gacggcatcc	gcaccctcgc	cgaccgcggc	gtcaccacct	tcgtcgaact	84060
cggccccgac	agcgtgctgt	ccgccatggc	ccaggagtc	gccccgaag	gcgcggcac	84120
catcccgtc	ctgcgcgcgg	accggcccca	ggaacaggcc	gtcctggccg	ccctctgcca	84180
cctccagggt	ctcggcgctg	aggccgactg	gtccgccacc	ttccgcggcc	tcgaccccgt	84240
ccgcgtcgac	ctgccgacct	acgccttcca	gcaccgctgg	ttctggcccg	ccgcccgaac	84300
cgcgcgcccc	gacgacgtcc	gcgcgcgcgg	cctgggcgcc	gccgaacacc	ccctcctcgg	84360
cgcgcgcgtg	caactccccg	acgacgacgg	cgcactcttc	accggccgcc	tctccctcgg	84420
caccaccccg	tggttggcgg	accacaccgt	cctgggcacc	gtcctgctcc	cgggcaccgc	84480
actggtggaa	ctgcgcgtcc	gcgcgggcga	cgagaccggc	agcggccacc	tcgaagaact	84540
caccctcgcc	gcgccccctga	ccctccccga	ggacggcgcc	accctcctcc	aggctccgct	84600
cggatccgcc	gacgacacgg	gccgcgcgac	cgtcaccgtc	cacgcccgcc	ccgacgacac	84660
cgcgcgaccg	acctggacgc	tgacgcgccac	cgggtgtgctc	gccaccacgc	caccggccgc	84720
cgcggcgctt	gacaccacgg	tctggccgcc	cgcgcgacgc	gaacccctca	ccaccgacga	84780
ctgctacgca	cacttcacca	cccaccgctt	cgcctacggc	cccgccttcc	agggcctcgg	84840
ggccgcctgg	cgcgcgcggg	acgtgctgta	cgcgcgaggtc	gccctgcccg	agtccgccac	84900
cgacgaagcg	gccgccttcg	gcctgcaccc	ggcgctcctg	gacgcgggcc	tgacgcgcgc	84960
gtcctcgcgc	gacgaccgcg	acaccggact	cccgttctcc	tgggaaggcg	tactctgca	85020
cgcctccggc	gccaccgcgc	tacgcgtccg	gtcgcgcccg	aacggcccca	acggcctgtc	85080
cgtcaccgcc	gccgaccgcg	ccggcaaccc	cgtgcgccacc	gtcaccgcgc	tgctcgcgcc	85140
ccccctggac	gccgagcagt	tgaccatcca	cagcgcctcg	accgcgcgac	cgtcttcca	85200
cctggactgg	accccggtcc	cgttccccga	caccgccaac	tccgcgcggc	cggccctcct	85260
cggcccggac	accgcgctgc	tcgccgacgc	cctcggcgac	ccggccgctg	cacgccacgc	85320
aaccctcgac	gacctcctgg	ccggggacac	caccccgccc	gccacggtec	tcgtccccct	85380
cggcgcccca	ctcgacggcg	acaccgcgca	gcacgcgcac	gccctcacc	gcagcgcgct	85440
gacctcgtc	cagcagtggt	tcgccaccga	ccgcctcgcc	gactcccgcc	tggtcttcgt	85500
caccacgga	gccgtcgcca	ccgacgacgc	gccccccacc	gacctggccg	ccgcgcgggt	85560
ctggggcctg	atccgctccg	cgcagaccga	gaaccccggc	accttcaccc	tcctcgacct	85620
cgacaccgag	cccgactcga	ccaccgcgct	cagccgcgcc	ctgacctcgc	acgaaccaca	85680
gtcctcctc	cgcgcggggc	gcgcccgcgc	cgcgcgcctc	accgcaccc	ccgccccac	85740
caccaccacc	cacacgcgct	ggtccgcgga	cggaaacgggtg	ttggtgacgg	gtggtacggg	85800



tggctctgggt	gggttgggtg	ccgggcatct	ggtgcggtcg	tgtgggggtgc	ggcatttgtt	85860
gttgaccagt	cgttctggtg	tgggtgctgc	gggtgcgggc	gggttgggtcg	cggagttgga	85920
gtcgttgggc	gcgcgggttg	tgggtgctgc	gtgtgatgtg	ggtgatggct	cggctgttgc	85980
ggagttgggt	gccgggtgtg	cggagtgcga	tccgttgtct	gcgggtgggtc	atgcggctgg	86040
tgtgttggat	gacgggtgtg	tgggttcgtt	gacgccggag	cggttgggtc	cgggtgttgc	86100
tccgaagggtg	gatgggtgct	ggaacctgca	tgaggcgacg	cgtgggtctg	atctggacgc	86160
gtttgttgct	ttctcgtctg	ttgcgggtgt	ggttcacggg	gtggctgggtg	ggttgcctgc	86220
ggcgggtaat	gccgtttttg	acgcgttgat	gggtgtgggg	atgacggcgg	ggctgacgga	86280
tgtgtcgttg	gcgtgggggtg	cttgggatca	gggtgtgggg	atgacggcgg	ggctgacgga	86340
gcgggatgtc	cgtcgtgctg	ctgagtcggg	tatgccgttg	ttgacgggtg	atcagggtgt	86400
ggcgttggtc	gatgcggcgt	tggcgacggg	gagtgccgcg	ttggtgccgg	tccgtctgga	86460
cctggccgca	ctgcgcaccc	ggggcgacat	cgcaccgtc	ctccgcggcc	tcgtccgcgc	86520
accgctgcgc	cgcaccgcgg	ccaccggcct	cgccaccggc	gcggacaccg	gcctcgtcca	86580
acggctcggc	cgactcgacc	acgccaacg	ccacgaggca	ctgctcgaca	tggctccgag	86640
cagcgcccg	ctcgtcctcg	gccacgccga	cggcaacgcc	atcgacgccg	aacgcgcctt	86700
ccgcgacctc	ggcttcgact	cgctcaccgc	ggtcgaactc	cgcaaccgtc	tgcgcacccg	86760
caccggcctg	cacctgtcgg	ccaccatggt	cttcgaccac	cccaccctgt	ccgcctcgc	86820
ggagcacctg	cgggacgagt	tgttcggcgc	ggtcgagagc	gaggtgaggg	tgccgggtcca	86880
ggcactgccg	ccgaccgccg	acgatcccat	cgtgggtggg	ggcatggcct	gccgtttccc	86940
cggtggtgtg	acctcgcccg	aggacctgtg	gcgcctggtc	gacgacggca	ccgacgccat	87000
caccaccttc	ccgaccaacc	gcggctggga	cctggacaac	ctctacgacc	cggaccccga	87060
gcacttcggc	acgtcgtaca	cccgcctccg	tggcttcctg	cacgaggcgg	gggagttcga	87120
cccggcggtc	ttcgggaatga	gcccgcgtga	ggcgctggca	accgactccc	aacagcgtct	87180
cctgctggaa	tccctcctggg	aggcgatcga	gcgggcccgg	atcgacccgc	tgaccctgcg	87240
cggcagcgcc	accggcgctc	tcgccggcgt	gagtacggg	gactacggga	gcatcctcgg	87300
cggcaaggag	ttcgagggct	tccaaggcca	gggaagtgcg	ggcagcggtg	cctcgggccg	87360
cgtctcctac	gccctcggct	tcgagggccc	ggcgcgtcac	gtggacacgg	cttgctcttc	87420
ctctctgggt	gccctgcact	gggcggcgca	ggcgttgccg	gcgggggagt	gctcgttggc	87480
gttgcccggt	ggtgtgacgg	tgatgtcgac	gccgagcacg	ttcgtggagt	tctcgcgcca	87540
gcgggggtctg	gcgcctgatg	gtcgttccaa	ggcgttcgcc	gaggcccgcg	acggcgctcg	87600
ctgggtccgag	ggcgtcggca	tcctcgtcct	ggagcgccag	tcggacgcgg	tgcgcaacgg	87660
ccacgagatc	ctcgccgtga	tcgcggctc	ggcgttaac	caggacgggtg	cgtccaacgg	87720
cctgaccgcg	cccaacggcc	cgtcccagca	gcgcgtcatc	cgtcaggcgt	tggccagtgg	87780
cggcctgtcc	acggccgacg	tggacgccgt	cgaggcgcac	ggcacgggta	cgacgctcgg	87840
tgaccgatc	gaggcccagg	cgctcctggc	cacctacggc	cgtgaccgcg	accccgagaa	87900
ccccctgtgg	ctgggctccc	tgaagtccaa	catcgggcac	acccaggcag	cggccgggtgt	87960
cgccgggtgtc	atcaagatgg	tcatggcgat	gcggcaacgg	gtgctgccgc	agaccctgca	88020
tgtcgacgcg	ccgtcctcgc	acgtcgattg	gagcgtcggc	gccgtcgaa	tgctcaccga	88080
gcagaccgcc	tggccggaga	ccggccgggt	ccgtcgcgcc	ggtgtctcct	ccttcggcat	88140
cagcggcacc	aacgcccacg	tcacgtgga	acagccggcg	ctcgtcgaaa	gcccggccgc	88200
ggagccgagc	ggacgcgaac	ccggcgctgt	tccgctgccg	ctgtccggaa	agtcccccca	88260
ggccctgcgc	gaccaggccg	cacgcctgct	ggccgggttg	gcggagcggc	ccgcgctgcg	88320
cccgtctgac	ctcggctact	cgctggcgac	gaccgcttcg	gcgttcgacc	accgggcggt	88380
ggtgctcgcc	accgaccgcg	ccgatgcggt	ccgcgcgtcg	acggcgctcg	ccgcggcgga	88440
cgcggatctc	tccgccgtcg	tcggcgacac	ccgcacgggt	cgtcacgcgg	tgctgttctc	88500
gggtcagggc	tcgcaacgcc	tgggcatggg	gcgtgagttg	tacgagcgtt	ttccggtctt	88560
cgccgaggct	ctcgatgtcg	cgatcgacca	cctggacgcc	gccttgcccg	cccaggccag	88620
tctgcgtgag	gtgatgtggg	gcgacgatgt	cgagctgctg	gacgagacgg	gttgagacga	88680
gccggctctg	ttcgccgtcg	aggtggccct	gttcggctcg	gtggagagtt	ggggtgtccg	88740
tccggacttc	gtggccgggtc	attccatcgg	tgagatcgcg	gcggcgcatg	tcgtcggggg	88800
gttctcgtcg	gaggacgcct	gcctgtctgg	ggccgcccgt	gcgacgctga	tgaggcgct	88860
gccgaccggc	ggcgcgatga	tcgcgatcca	ggccgcccag	gacgaagtca	cccagcacct	88920
gactgacgac	gtctcgatcg	ccgccgtcaa	cggcccagacc	tctgtgggtcg	tctccggggc	88980
cgagagcgct	gcccgcacgg	tggccgaccg	gctcgcggag	aacggccgca	agacgacccg	89040
gctgcgggtc	tcgcatgcgt	tccactcgcc	gttgatggat	ccgatgctgg	cggagttccg	89100
tgcggtggcc	gagggcctgt	cctacgccac	cccgaccctc	cccgtcgtct	cgaacctgac	89160
gggcccggctg	gccacggccg	atgacctctg	ctcgcccgag	tactgggcgc	gccacgtccg	89220
cgaggcgggtc	cgcttcgccg	acggcgctcag	caccctggag	aacgagggcg	tcaccaggtt	89280
cctggaactg	ggaccggacg	gcgtgctgtc	cgccatggcc	cagcagtcgc	tcaccggcga	89340
cgcgccacc	gtcccggccc	tccgcaagga	ccgcgacgag	gagacgtccg	cgctcaccgc	89400

cctcgcccac	ctccacacgg	caggtctccg	cgtcgactgg	gcggcggttct	tcgctcggcag	89460
cggcgccacc	cgcgtggacc	tgccgaccta	cgccttccag	cacgccacct	actggcccac	89520
cggcaccttg	cccaccgcgc	acgccgcggc	cgtcggcctc	accgccgccg	agcaccgcgt	89580
gctgaacggt	tcggtcgaac	tcgccgaagg	cgaaggggtg	ttgttcaccg	gacggctgtc	89640
actgcagtca	catccgtggc	tggccgacca	cgccgtcatg	ggacaggtcc	tgctgccccg	89700
caccgcactg	ctggaactgg	cgttccgggc	cggcgacgag	gccggttgcg	accgcgtcga	89760
ggaactgacg	ctcgccgcac	cgctcgtcct	gcccgcgcgc	ggtgcggtac	agaccaggt	89820
ccgggtcggc	gtcgccgacg	acaccggccg	ccgtaccgtc	accgtccact	cccggccccga	89880
gcacgcgacc	gacgtgtcgt	ggaccagca	cgcgaccggc	accctgacca	tgggctccgc	89940
cccggccgac	accggtttcg	acgccactgc	ctggccgccc	gccgacgccg	aaccctcgc	90000
caccgacgac	tgctacgcgc	gcttcacgac	gctcggttcc	gcctacgggc	cggctctcca	90060
gggcctgcgg	gccgcctggc	gcgccggtga	cgtgctgtac	gccgaggtgg	ccctggcgga	90120
gtccaccggc	gacgaggcga	ccgccttcgg	tctgcacccc	gcaactgctc	acgccgcct	90180
gcacgcctcc	ctcgtcgcgc	acgagggcga	ggagagcaac	ggcggactgc	cgttctcctg	90240
ggagggcgcg	accctctacg	cgaccggcgc	caccgcgctg	cgcgtccggc	tgaccccgac	90300
gggcaccgac	ggccgttcgg	tggccatcgc	cgtggccgac	accgccggtc	gtccggtcgc	90360
cgccatcgac	aacctcgtct	cgcgcggggt	ctccggcgac	cagttgaccg	gcgcccgggg	90420
actggcccgc	gacgccctgt	tcacctgga	ctggaacccc	gtaccggaga	acctcgtacc	90480
ggagaacctt	gtaccggaga	acaccggcgg	gggccacgcc	caggaccagg	acggccggcc	90540
cgcgcgggcc	accgtcgcgc	tggtcggcgc	ggacggcacc	gcgatcgccg	ccgacctgac	90600
gcgcgcgggc	atccacacca	ccctccaccc	cgacctcacc	accctcgcca	cgaccgacgc	90660
cgcggttccg	aagacgggtc	tcattccccct	caccggaaac	ggaaccggaa	ccggcaccgg	90720
gactgagtcg	acggacggaa	tcgggacggg	ggccgcgcgag	tcggacgcgt	ccgccccctc	90780
cccggccgag	gtcgcccaca	ccctgtccac	gcgcgcactc	gccctcgtcc	aggagtggac	90840
cgcacaggag	cgttccgcgc	gctcccgctt	ggcgttcgtc	acgaccgggg	cgacggccgc	90900
cggcggtacc	gacgtcatgg	acgtggccgc	cgcgcgggtc	tggggcctgg	tccgatccgc	90960
ccagtccgaa	gccccggaca	ccttcgtcct	gatcgaccgt	gaccccggcc	cggccggcac	91020
gcacgaccgc	acagccgcgc	ccgaacgggg	ccaactgctc	ctacgggcac	tgcacaccga	91080
cgaaccgcag	ctcgccctgc	gtgacggcgg	cgtgctcgcc	gcccgccctg	cccgtctcga	91140
caccgcggcc	gcgctcacc	cgcgcggcga	ccgggcctgg	cggctcgaca	gcacggccaa	91200
gggcagcctc	aacggcctcg	ccctgacccc	gtatccggcg	gcaactggcg	cgctcaccgg	91260
ccacgaggtg	cgggtcgagg	tgcgtgccgc	gggcctgaac	ttcgtgacg	tgctcaacgc	91320
gttgggggatg	tatccgggtg	atgatgtcgg	atcgttcggg	tcggaggcgg	ccggtgtggt	91380
cgtcgaggtc	ggaccggagg	tgaccggcct	ggcccccggc	gaccaggtca	tgggcatgat	91440
caccggcagc	ttcggtcgc	tcgccgtgga	cgacgcgcgc	cgcctcgccc	gcctgcccga	91500
ggactggtcc	tgggagacgg	gtgcgtcggt	gocgttggtg	ttcctcaccg	cgtactacgc	91560
cctgaaggag	ttgggtggtc	tgccggcggg	ggagaagggtg	ctggtgcatg	ccggtgccgg	91620
tggtgtcggt	atggcggcga	tcagatcgcc	ccggcatgtc	ggtgccgagg	tggtcgccac	91680
ggccagttag	ggcaagtggg	acgtgttgcg	ctcgtcgggc	gtggccgacg	accacatcgc	91740
ctcctcccgc	accctcgact	tcgaggcggc	cctcgccgaa	gtcgccggcg	accgcggcct	91800
ggacgtcgtc	ctcaactccc	tcgcgggtga	ccttcgcgac	gcctcgatgc	ggttgcctcg	91860
cgcggcgccg	cgtgtcctgg	agatgggcaa	gaccgacatc	cgcgcgcggg	actccgttcc	91920
cgcggccctc	tcctaccagt	ccttcgacct	cgcctgggtg	gtgccggaaa	ccatcggcac	91980
catgctggcc	gagctgatgg	acctcttccg	caccggcgca	ctgcggccac	tcccggctcg	92040
cacctgggac	gtccggcacc	ccaaggacgc	gttcgcgttc	atgagcatgg	ccaagcacat	92100
cggcaagatc	gtgctcacc	tgccccgctc	ctggaagccc	gagggaacgg	tgttggtgac	92160
gggtggtacg	ggtggtctgg	gtgggttggt	ggcccgccat	ctggtgcggt	cgtgtggggg	92220
gcggcatttg	ttggtgacca	gtcgttctgg	tgtgggtgct	gcgggtgcgg	ccgggttggt	92280
cgcggagttg	gagtcgttgg	gcgcgcgggt	tgtgggttgc	gcgtgtgatg	tgggtgatgg	92340
ctcggctgtt	gcggagttgg	ttgccgggtg	gtcggagtcg	tatccgttgt	ctgcggtggt	92400
gcatgcggct	ggtgtgttgg	atgacggtgt	ggtgggttcg	ttgacgcggg	agcggttggc	92460
tgcggtgttg	cgtccgaagg	tggatggtgc	gtggaacctg	catgaggcga	cgcgtggtct	92520
ggatctggac	gcgtttgttg	tcttctcgtc	tgttgccggg	gtgttcgggg	gtgcgggtca	92580
ggccaactat	gcggcgggta	atgcgttttt	ggacgcgttg	atggttcac	gggtggtcgg	92640
tgggttgcc	ggtgtgtcgt	tggcgtgggg	tgcttgggat	cagggtgtgg	ggatgacggc	92700
ggggctgacg	gagcgggatg	tcgctcgtgc	tgctgagtcg	ggtatgccgt	tgttgacggg	92760
tgatcagggt	gtggcggtgt	tcgatgcggc	ggtggcgacg	gggagtgcgg	cgttggtgcc	92820
ggtccgtctg	gacctggccg	cactgcgcac	ccggggcgac	atcgcaccgc	tcctccgcgg	92880
cctggtcaag	gcgcccatcc	gcgcgcgcgc	cgcaccaca	cccggcgaca	ccggactcgc	92940
cgcgcagctc	accgggctcc	agcgcgcgcga	gcgacgggac	accctcctcg	cgtcgtcgcg	93000



cgaccaggcc	gcatggtcc	tcggccacac	ctcgggagac	ggcgctgacc	cgtcccgcgc	93060
cttcgcgcac	ctcggcttcg	actcgctcac	cgcggtcgaa	ctccgcaacc	gcatcggcgc	93120
ggccaccggc	ctgcggttac	cggccacggc	cgtcttcgac	tacccacccg	ccgatgccct	93180
cgccgcacac	ctgctcaccg	aactgctcgg	ccccgacgcc	gagtcggacc	ccgacgagcc	93240
cgggcagccc	accgcgggac	cgaccgacga	ccccatcgtc	atcatcggca	tgagctgccg	93300
cttccccggc	gacatcggct	cgccggagga	cctgtggcgc	ctgctcggcg	acggcgccga	93360
cgctcgtcacc	gacttcccga	ccaaccgcgg	ctgggacctg	gacaacctct	acgaccccga	93420
ccccgcgcac	gccggcacct	cgtacgcccg	caccggcggg	ttcctgcacg	acgcccgcga	93480
cttcgacgcc	gacttcttcg	gcatgagccc	ccgcgaggcc	atggccacgg	actcccagca	93540
gcgcctgctg	ttggagtcct	cgtgggaggc	gatcgagcgg	gccggcatcg	acccgctgac	93600
cctgcgcgac	agccgcaccg	gcgtcttcgc	cggcgtcatg	tacagcggct	acggcacccg	93660
cctcgacggc	gccgaattcg	aaggcttcca	ggggcagggc	agcgcactga	gcgtggcctc	93720
cgggcgggtc	tcctacacct	tcggcttcga	aggcccggcc	atgacggctg	acaccgcctg	93780
ctcctcctcg	ctggtcgccc	tgcacctcgc	cgcacaggca	ctccgcggcg	gtgagtgac	93840
cctcgccctc	gccggtggtg	tcaccgtgat	gtccatcccg	gacaccttca	tcgagttctc	93900
ccggcagcgc	ggactggccc	ccgacggccg	ctccaagccg	ttctccgagt	ccgccgacgg	93960
cgtcggctgg	tccgagggcg	tcggaatgct	gtcctgggag	cgccagtcgg	acgccgtgcg	94020
caacggccac	cagatcctgg	ccgtggtgcg	cggtcggcg	gtcaaccagg	acggtgcgtc	94080
caacggcctg	accgcgccca	acggcccgtc	ccagcagcgg	gtgatccgtc	aggcgttggc	94140
cagcggcggc	ctgtccacgg	ccgacgtgga	cgccgtcgag	gcgcacggca	cgggcaccac	94200
gctcggtgac	ccgatcgagg	cccaggccct	cctggccacc	tacggcccg	accgcgaccc	94260
cgagaacccg	ctgctgctcg	gttcgatcaa	gtccaacctc	ggccacaccc	aggcagccgc	94320
cgggtgcgcc	ggcgtcatca	agatggtcat	ggcgatgcgg	cacggcgtgc	tgccccgcag	94380
cctgaacatc	accgagccgt	cctcgcacgt	cgatggagc	gccggcgccg	tcgaactgct	94440
caccgagcag	accgcctggc	cggagaccgt	cggggcccg	cgcgccggta	tctcctcctt	94500
cggcatcagc	ggcaccaacg	cccacgtcat	cctggagcag	ccggaggccg	cgcggcactc	94560
ggcgccggaa	gaagccgaca	cggcggaggc	agccgccaa	gcgccggcca	ccgcgcacct	94620
gcccgtaatg	ccgtgggcac	tgtccggcaa	gacgcggag	gccctgcgtg	cccaggccgc	94680
acgcctcctc	gcccacctcc	agcagcgcgc	cgaactcgca	cccgccgaca	tcgccctgtc	94740
cctcgccacc	cagcgtctcc	agttcaccca	ccgggcagtc	gtcctgagca	ccgaccgtga	94800
cgaggcgacc	cgcgcgctgt	ccgccctcgc	caccaccgcc	gcgtccgacc	cctcggccct	94860
caccggcacg	gtcaccatgg	gacgttgccg	ggtgctgttc	tcgggtcagg	gctcgcaacg	94920
tctgggcatg	gggcgtgagt	tgtacgagcg	tttcccggtc	ttcgccgagg	ctctcgatgt	94980
cgtgatcgat	cacctggacg	ccgccttgcc	cgcccaggcc	ggtttgctg	aggtgatgtg	95040
gggcgacgat	gtcgagttgc	tgaacgagac	gggttgagcc	cagcccgcgc	tcttcgccat	95100
cgaggtggcg	ctgtttcggc	tggtggagag	ttgggggtgc	cgtccggact	tcgtggccgg	95160
tcattccatc	ggtgagatcg	cggcgccgca	tgtcgtcggg	gtgttctcgt	tggaggacgc	95220
gtgccgtctg	gtggccgcgc	gggcgacgct	gatgcaggcg	ttgccggccg	gtggcgcgat	95280
gatcgcggtc	caggcgaccg	aggacgaagt	catcccgcac	ctgaccgacg	aggtggcgat	95340
cgcggccgtc	aacggcccga	cctccgtggt	gatctcgggc	gcagaagagg	ccacgcagac	95400
cgtggcacaa	cacttcgccg	accagggggc	cgggacgacc	gcgtgcggg	tctcgcctgc	95460
gttccactcg	ccgctgatga	tgctggcgga	gttccgtgcg	gtggccgagg	gcctgtccta	95520
cgccaccccg	accttccccg	tcgtctcgaa	tctgacgggc	caggtggcca	cggccgacga	95580
actctgctcg	gccagtagct	gggtgcgcca	cgctcgtgag	cgggtccgct	tcgccgacgg	95640
tgtgacggcc	ctcgaaagcc	agggcggtcg	gaccttcctg	gaactcggcc	cggacggcgt	95700
cctcgccgcc	atggccaggg	aaaccgtcgc	cgacgacacg	gtcaccgtcc	ccgtcctccg	95760
caggaacatg	cccaggaac	ggacctgct	caccgcactc	ggccggctcc	acaccaccgg	95820
aaccccgatc	gactgggcgc	ccctcctggc	cccgaccggc	gcccgcgccg	tggacctgcc	95880
gacatacgcg	ttccaacacc	gtcccttctg	gccctccggc	ccccgcgaca	ccgcggatgc	95940
cgccgcgcgc	ggcatcgccg	gcgcgagcca	cccgctcctc	aacggcatcg	tcgaactcgc	96000
cgacgaagag	ggcctgttgt	tcaccggacg	gctgtcactg	cagtcgcata	cgtggctggc	96060
cgaccacgcc	gtcatgggac	aggctcctgt	gcccggcacc	gcactgctgg	aactcgccct	96120
gcgcgcgggc	gacgaggtcg	gctgtgacca	cgtcgaggaa	ctgacgctcg	ccgcaccgct	96180
cgtcctgccc	gagcgcggcg	cggtacagac	ccaggtccgg	gtcggcgctg	ccgacaccac	96240
cgggcgcgcg	accgtcacga	tccactcgcg	tcccgcacgc	gccacgacca	ccgacagtga	96300
caccacaccc	ggcaccgaca	ccccgtggac	ccaacacgcc	accggcgctc	tcgtcgccgg	96360
cctgccggcg	acggcaaccg	tcccgttcga	tgccaccgtg	tggccgccgg	cgcacgccga	96420
acccgttgac	ctggcgggact	tctacgcgtc	ccgggcgggc	gaaggattcg	gctacggggc	96480
cgttttccag	ggcctgcgag	cggcctggcg	ccgcgacggc	gaggtgttcg	ccgatgtcgc	96540
actgccggag	gccggccgta	ccgaagccga	ggcgtacggg	ctgcatccgg	cactgctcga	96600

cgccggactg	cacgcagcct	ggctcgtcgc	cccggacggg	gagccacac	ggacgggcag	96660
cgtgccgttc	tcgtagcgcg	gcgtcttctt	ggcgcgttcc	ggtgcctcct	cggtcgcggt	96720
ccgactcggc	cgcgactccg	acggaacgct	gagcctggcc	atcgccgaca	ccaccggtgc	96780
accggtcgcg	tccgtacagg	ccctctccat	gcgcaccgtc	tccgtgacgg	ccctgagcgc	96840
caccgcgggc	ctcgcccgcg	acgcgtgtgt	ccgcctggac	tgggcctcgg	ccccggagcc	96900
ggcgtgccag	ccggacgaca	cggtagccgt	gatcccggcg	gtcgcggtcg	tccgtacgga	96960
aacctccgaa	ctcacctccg	agctcaccgc	ggccttgcgt	gcggcggggc	ccgacgtcga	97020
cgtccgcacg	accctgtcga	ccgatgaacc	cgcgcgcgcg	ctgatcgcgc	tgcgcgtcgt	97080
cgcctccgac	cagaccggca	ccgcagaagc	ggcaccgttc	ccggcggccg	tgcacgacct	97140
cacccgccga	gccctggccc	tccgtacagac	ccgcctgcaa	gagcagcact	tccgcggacac	97200
gaagttcgtc	tccgtgaccc	gtggtgacgc	ggcggggcgt	gatgtggctg	ctgctgcggt	97260
gtggggtctg	gtgcgttcgg	cgcagtcgga	gaatccgggt	tgttttgctc	tggctgatct	97320
ggatccggat	ggtgcgggtg	gtgcggctgc	gctcgtcgtc	gcgttggtca	gtggtgagcc	97380
gcagcttgcg	gtgcgcgggt	atgtgttgcg	ggcgcgcgtc	ctggtgcggc	ggccgctcac	97440
cagagtcggg	gcgggtgctg	atggcaccgg	ggatggcgct	gggggtggct	ctggtgtgtc	97500
gttctcgggt	gaggggtgcg	tccgtgtcac	tgggtgtacg	ggtggtctgg	gtgcggtgtt	97560
ggcgcgtcat	ctggtggccg	agtatggggg	gcgggatctg	ctgttggtca	gtcgcagtggt	97620
tgaacgtgcc	gtgggtgctg	gggagttggt	ggcggagctt	gcgggtgtgg	gtgcgcgggt	97680
gcgggtgggt	gcgtgtgatg	tgaccgatcg	tgcgcgggtg	gtggagttgg	tggcggggca	97740
tgcggtgtcc	gcggtgggtc	atgcggctgg	tgtgctggat	gacggcatgg	tgggtgcgtt	97800
gaccggggag	cggttgtccg	cgggtgctgc	gccgaagggt	gatgctgtct	ggcatctgca	97860
tgaggcgacc	cgcggcctgg	atctggacgc	gttcgtcgtc	ttctctctct	tccgcggggg	97920
cttcggcagt	cccggccagg	ccaactacgc	ggcgcgaac	gccttctctg	acgcgctgat	97980
gacgcggcgc	cgggcggagg	gactgcccgg	cctgtcactc	gcatggggac	cgtggtcgct	98040
gaccgatggc	acatcgggca	tgctcgcgga	cgcgcaggcc	gatcgcctga	ccgctgcggg	98100
agtgccaccg	ctgaccgcgg	agcaaggact	ggcactgttc	gacgcggccc	tggcgaccgg	98160
tgacgccacc	tgcgtcccgg	tccgcctgga	cctgtcggca	ctgcgtgccc	aggggtgaggt	98220
gccgcccttg	ctgcggtccc	tgatccgagg	ccgctcgcgc	cgcgcgcgcg	ccgcggaatc	98280
ggcaaccgcc	accgggctgc	gggaacggct	cgtcggactg	aaccgggtcg	agcgacagga	98340
agtcctcctg	gacctcgtgc	gcggccaggt	ggccttggtc	ctcgccacag	ccgacgccga	98400
cgacgttcat	ccggcacgtg	ccttcaggga	gttgggcttc	gactccctca	cctcggtgga	98460
actgcgcaac	cgcctcaaca	ccgtcaccgg	tctgcggctc	cccgcgacca	tgggtgttcga	98520
ctatccgacc	gtcgaagtgc	tggctctcta	cgtcctggac	gagttgctgg	gcacggatgc	98580
cagagtgggc	accgtgcagc	cggccgcggg	tgcggtggcg	gacgatccga	tgcgtcatcg	98640
gggcatggcc	tgcgcgtacc	ccggcggggg	cgcctctccg	gacgacctgt	ggaggctcgt	98700
cacggacggc	gtcgcacgcg	tgtccccctt	cccgaaccaac	cgtggttggg	acgtcgaatc	98760
cctctatcac	ccggaccccc	accatctcgg	tacctcctac	acgcgctcgg	gtgggttcct	98820
gcatgaggcg	ggggagttcg	atccgggggt	cttcgggatg	agtcgcgcgg	aggcgttggc	98880
gaccgattcc	cagcagcggg	tgttgttgga	gtcgtcgtgg	gaggcgatcg	agcgggcccg	98940
tattgatccg	gtgagtttgc	ggggtagtcg	gacgggtgtg	ttcgcggggg	tgatgtacag	99000
cgattacagc	gcgatgttgg	cagatccgga	gttcgagggt	ttccagggca	gtgggagttc	99060
gccgagtttg	gcgtcgggtc	gggtggccta	cacgttgggg	ttggaaggcc	cggcggtgac	99120
ggtggatacg	gcgtgttcgt	cgtcgttggt	ggcgatgcac	tgggcgatgc	aggcgttgcg	99180
tagtggtag	tgtgggttgg	cgttggccgg	tgggtgtacg	gtgatgtcga	cgcctgcggg	99240
gtttgtggac	tttgctcggc	agcgggggtt	gtcgcgggat	ggtcgggtgca	aggcgtttgc	99300
ggatgcggcc	gatggtgtgg	gctgggtccga	gggcgtcggc	gtgttggtcc	tggagcggca	99360
gtcggacgcg	gtgcgcaatg	gtcacgagat	tttggctgtg	gtgcgggggt	cggcgggtcaa	99420
ccaggatggt	gcgtccaatg	gtttgacggc	gcccaatggt	ccgtcgcagc	agcgggtgat	99480
ccggcaggcg	ttggccagtg	gtggtctgac	ggcgggtgac	gtggatgtgg	tggaggcgca	99540
tggtagcggg	acgacgctcg	gtgatccgat	cgaggcgag	gcgttggttg	cgacgtatgg	99600
gcgggatcgt	gagcctgagc	ggccgttggt	gttgggttcg	gtgaagtcca	atctggggca	99660
tacgcaggct	gctgcgggtg	tggcgggctg	tatcaagatg	gtgttgggca	tgcggcatgg	99720
tgtggtgccc	cggacgttgc	atgtggatgc	gccttcttcg	catgtggact	ggtccgaggg	99780
tgcggtggag	ctgctcagtg	agcaggcggc	ttggccggag	acgggtcggg	tgcggcgggc	99840
gggtgtctcc	tccttcggca	tcagcggcac	caacgcgcac	gtgattctgg	agcgtccgga	99900
ggccgcgcgg	cgtccggtga	tggagacgaa	caccgtggag	ccgtccacgg	tgccgtgggt	99960
gctttccggc	aagacaccgg	aggccctgcg	cgcgcaggcg	gcgaaactgt	tgtcgtcgat	100020
cgaggaaacgc	ccggagcttc	gcctcgtcga	cgtgggaatg	tccctgggtca	ccggccgctc	100080
gacctttgaa	caccgtgcgg	tggtagctgc	cgcgcaccgt	gccgacgccg	cccgcgcatt	100140
gtcggcgatt	gctgccgacg	aggccgatgc	tgcgcgtgcc	accggcccg	tccgcgcggg	100200

tcgtcacgcg	gtgctgttct	cggttcaggg	tgtctaacgt	ctgggcacgt	ggcgtgagtt	100260
gtacgagcgt	ttcccgggtct	tcgccgaggg	tctcgatgtc	gtggtcgacc	acctggacgc	100320
cgccttgccc	gcccaggccg	gtttgcgtga	ggatgatgtg	ggcgacgatg	cggagttgct	100380
gaacgagacg	ggttggacgc	agccggctct	gttcgccatc	gaggtggcgc	tgttcgggct	100440
gggtggagagt	tgggggtgtcc	gtccggactt	cgtggccggg	cattccatcg	gtgagatcgc	100500
ggcggcgcat	gtccgcgggg	tgttctcgct	ggaggacgcc	tgccgcctgg	tggccgccc	100560
tgcgacgctg	atgcaggcgt	tgccggccgg	cggatcgatg	atcgcggtcc	aggcgaccga	100620
ggacgaagtc	accccgcatc	tgaccgacga	cgtagcgatc	gccgccatca	acgggcccga	100680
cgcactggtc	gtatcgggtg	tggaggatgc	cgcgcgtcag	atcgggggcg	ggttcgcggc	100740
cgagggggcgt	cgcacgaccc	gactccatgt	gtcgcgatgc	ttccactcgc	cgttgatgga	100800
tccgatgctg	gcggagtcc	gtgtgggtgg	ggagggcctg	tcctatgctg	ctccgtccct	100860
ccccgtcgtc	tcgaatctga	cgggccaggt	ggccacggcc	gacgaactgt	gctcggccga	100920
gtactgggtg	cgccacgtcc	gcgaggcggg	ccgcttcgcc	gacggggatg	cggccctcga	100980
agccgagggc	gtgcggacct	tcctggaact	cggcccggac	ggcgtcctcg	ccgccatggc	101040
aggagcctcg	ctcaccgaat	ccctccctcg	ggtagcgctg	ctccgtaagg	accggccgga	101100
ggaaccggcg	gcactcgccg	ccctggccca	ggtgcacatc	gccggcgcg	gcgtcgattg	101160
gcccgtgctc	ttcgtcggtg	tgggtgcggg	gcgggtggag	ttgccgacgt	atgcgttcca	101220
gcgtgggtgg	ttctggccgg	ttgggtcggg	tgggtgtggg	ggatgatgtg	gtgctgtggg	101280
gcttgggtct	gcggggcacc	cgttgttggg	tgtgcgggtg	gagttggctg	cgggtgcggg	101340
gggtggtgtg	acgggtcgct	tgtcgttgtc	gtcgcgatgg	tgggtggctg	atcatgcggg	101400
gatggggcgg	gtgtttgttc	ctgggtacgg	gttgcgtggg	atgggtgatg	gtgctgggtg	101460
tgaggtgggg	tgtggtcgct	ttgaggagct	gacgttggcg	gcgcggttgg	tgttgccctga	101520
gcgtgggtgg	gtgcgggttc	agggttgcgt	ggatgctcct	gatgctgcgg	gtcgtcgctg	101580
tgtgggggtg	tattcggtgc	ctgatgggtg	gggtcaggcg	gtgtgggtcg	agcatgcgtg	101640
cgggtgtgtg	gcctctgggt	tgggtgacca	gggtcgggtg	ttcgggtgac	gtgggtgtgtg	101700
gccgcgcgag	gggtgcgggt	cgggtggatg	tgagggtcgc	tacgagctgt	ttgcggatgc	101760
tgggttcggg	tatggcccg	tgttcagggg	gttgcggtgc	gtgtggcgct	gcggcgagga	101820
actcttcgca	gaggtcgccc	tgtcggacga	gggttgcgtg	agcgctgata	cggcgaccgg	101880
tttcgggttg	caccgcggct	tgtcggatgc	ctcgtctccat	gcctcgcttc	tctcctccct	101940
tgaaggtcaa	tccgccgatg	gtgggcctgc	gttgccgttc	gcgtgggagg	gtgtttccct	102000
cttcgcctcg	gggtgcgacg	ctttgcgcgt	gcgggttggc	ccggcgggtg	agcatgcggg	102060
gtcggtgacc	gcgggtggatc	cgcacgggtg	gccgggtgatt	tccatcgacg	cgttcctgtac	102120
ccgtcgccct	accctcgatg	agggtcaacgc	atctcacacc	cagctgagcg	atgcgtctt	102180
cggcgtccaa	tggaccacgg	tcccagagcac	cccggccgcc	gaccaccctg	cggtagccat	102240
catcggaacc	gaccacctgg	ggctcgccga	agcgtcagc	agttcctctg	ctggcgcgac	102300
cacgaccacg	accgcggccg	cgtacgagag	ccttgacgcg	ctgatcgcg	cggggcccga	102360
agtgtccgtc	cctgacgtca	cactcatcgg	tctcaccacc	gaggacgcca	tcgctcagta	102420
cgtgaacgac	cacgacgcca	cgggtggccg	ccaaggcacc	atcggagccg	gcgcggcgcc	102480
cgtagacgcg	gctcgccgcc	tcaccgcgga	ggccctgcgc	acgatccagg	catggttggc	102540
gcacgagcgt	ctggcggcaa	ggcgccctgg	cttcgtgacc	cgcggcgccg	ccgacgggca	102600
ggatgtgcgc	gcggcgcccg	tccagggcct	gggtgcgtcc	gccagaccg	agaaccgggg	102660
caccttcggc	ctcctcgacc	tgcacgggac	caggcgctcg	accgcgggtc	tcggcgaggc	102720
tctcacctcc	gacgaaccgc	aactgcttct	gcgcgacgga	cacctgcacg	ccgccaggct	102780
gacccgcctg	gcgtcgcccg	ccgacacggc	cgtgccacg	gagtggaaac	cggacggcac	102840
gggtgctgatc	accgggtggtg	cggcggggct	cggcgcgacg	ttcgacggc	acctcgtcga	102900
caggtagcgg	gtccgcaatc	tctgtctcgt	cagccggcgc	ggccccgatg	ccccgggaac	102960
cacggagtgtg	gtcgcgcgag	tgacggcgca	cgggtgcgag	gtggccgtgc	aggcatgtga	103020
cgtggccgat	ggcgatgcgg	tggcggcgtt	gggtgcgggc	gtgccggatg	agcaccgcct	103080
gagggcggtc	gtccacacgg	ccgggtgtgtt	ggacgacgga	gtgatcggct	cgtcaccga	103140
ggagcgggctc	gccaccgtcc	tgcggcccaa	ggcggatgcc	gcctggcatc	tgcacgaagc	103200
gacccgcggc	ctggacctgg	acgcgttcgt	cgtcttctcc	tccgtcgacg	gggtcttcgg	103260
cggcgccggc	caggccaact	acgcgcgggc	caacgccttc	ctggatgcgc	tgatggccca	103320
gcgcggggca	gcgggcctgc	ccggactctc	gttggcctgg	gggcccgtgg	accagaccgg	103380
cggaatgacg	ggcatgctgt	cggacgccga	ggccgaccgc	ctcgcccgtc	ccggcatccc	103440
gccgtctctc	gcggagcagg	gcctcgccct	cttcgacgcg	gcactcgctc	ttgccggaac	103500
cagcacgcca	gacagggcag	ccggcagcgc	cgcgcgcagc	acgtcgggaa	ccggcgacac	103560
gatcgccatc	ccggccgcgg	ccctcgctgc	accggctcgg	ctcgacctgg	cggcgctggc	103620
cgcgcagggt	gaggtccccg	cgattctgcg	cggactgggtg	cggaccgcga	cccgccttac	103680
ggcggccggc	ggttccgtca	ccgtggccgg	actcgtcaac	cgcctgtccg	ggctcaccgc	103740
cgacgagcgg	cgccaggaaac	ttctcgaact	ggtcgcgact	caggcagccc	tcgtcctcgg	103800

gcacgcccgat	ccggcgctccg	tggactccac	cgcacagttc	cgtgacctcg	gcttcgactc	103860
gctgaccgcc	gtcgagctgc	gcaaccggct	gagcacggcc	accggcctgc	gcctgaccgc	103920
aaccctggtc	ttcgactacc	cgaacaccga	tgccctcgcg	gagcacctgc	gggacgagtt	103980
gttcggcgcg	gtcgagagcg	aggtgcgggt	gccgggtccag	gcaactgccg	cgaccgccga	104040
cgatcccac	gtgggtgggtg	gcatggcctg	ccgtttcccc	gggtggtgtga	cctcgcccga	104100
ggacctgtgg	cgcctgggtc	acgcccggac	cgacgccatc	accaccttcc	cgaccaaccg	104160
cggctgggac	ctcgaatcgc	tctacgacct	ggaccccgc	cacctcgga	cctcctacac	104220
ccgctccggt	ggcttcctgc	acgaggcggt	ggagttcgac	ccggcgttct	tcggaatgag	104280
cccgcgtgag	gcgctggcaa	ccgactccca	acagcgctct	ctgctggaat	cctcctggga	104340
ggcgatcgag	cgggcccggca	tcgacccgct	gacctgctgc	ggcagcgcca	ccggcgctct	104400
cgccggcggtg	atgtacagcg	actacgggag	catcctcggc	ggcaaggagt	tcgagggtct	104460
ccaaggccag	ggaagtgcgg	gcagcggtgc	ctcggggcgc	gtctcctaca	ccctcggtct	104520
cgaagggtccc	gccgtcaccg	tggataccgc	ctgctcctcc	tccttggtcg	ccctgcacct	104580
ggcagcccag	gcccttcggg	cgggtgagtg	cacgctcgcg	ctcgccgggtg	gtgtgacggt	104640
gatgtccacg	ccaggcacgt	tcgtggagtt	ctcgccggcag	cgggggtctgg	cgcctgatgg	104700
tcgttccaa	gcgttcgccc	aggccgcgga	cggcgctcggc	tggtccgagg	gcgtcggcac	104760
cctcgctcctg	gagcgccagt	cggacgccgt	gcgcaaccggc	cacgagatcc	tcgccgtgat	104820
ccgcggctcg	gcggtcaacc	aggacggtgc	gtccaaccggc	ctgaccgcgc	ccaaccggccc	104880
gtcccagcag	cgcgtcatcc	gtcaggcggt	ggccagtggc	ggcctgtcca	cggccgacgt	104940
ggacgccgtc	gaggcgacg	gcacgggtac	gacgctcggt	gacccgatcg	aggcccaggc	105000
gtcctggcc	acctacggcc	gcgaccgcga	ccccgagaac	ccgctgctgc	tcggttcgat	105060
caagtccaac	ctcggccaca	cccaggcagc	ggccgggtgc	gccggtgtca	tcaagatggt	105120
catggcgatg	cggcacggcg	tgctgcccga	gacctgcac	gtcgacgcgc	cgtcctcgca	105180
cgtcgattgg	agcgtcgcg	ccgtcgaaact	gctcaccgag	cagaccgtct	ggcccgagac	105240
cggccgggtc	cgtcgccgcg	gtgtctcctc	cttcggcatc	agcggcacca	acgcccacgt	105300
catcctggaa	cagccggagg	ccgtgcagcg	cctggcaccg	ggagcagcag	agaccgtgga	105360
gccggtcgcc	atcaagccgt	cggcggaacc	gtccctggtg	ccgtgggccc	tgtccggcaa	105420
gtcaccgcag	gccctgcgcg	cccaggccgc	acgcctccgt	gacttcctgg	cggaacggcc	105480
cgaaccgcgc	tcgatcgaca	tcggccactc	actggccgtc	acacgctcgc	agttcgacca	105540
ccgcgcgatc	gtgctggtcg	acgatgcgaa	ggccccggcg	gacagcctgg	ccgccctcgc	105600
ggccctggcc	tcgggtgtgg	ccgatcccgc	cgtcgtctcc	gacgcggtat	cgaccggcgg	105660
ttcggcagtg	ctgttcacag	gtcagggtgc	tcaacgtctg	ggcatggggc	gtgagttgta	105720
cggccgtttc	ccggtcttcg	ccgaggctct	cgatgtcgtg	gtcgaccacc	tggaacgcgc	105780
cttgcccggc	caggccggtt	tgcgtgaggt	gatgtggggc	gacgatgtcg	agttgctgaa	105840
cgagacgggt	tggaaccagc	ccgcgctctt	cgccgtcgag	gtggcgctgt	tcgggctggt	105900
ggagcgttgg	ggtgtccgtc	cggacttcgt	ggccggtcat	tccatcggtg	agatcgcggc	105960
ggcgcatgtc	gccgggggtg	tctcgctgga	ggacgcctgc	cgtctggtgg	ccgcccgctg	106020
gacgctgatg	caggcgctgc	cgaccggcgg	cgcgatgac	gcggtccagg	ccaccgagga	106080
cgaagtgacc	ccgcacctga	ccgacgaggt	ggcgatcgcg	gccgtcaacg	gcccgaacct	106140
cgtggtgatc	tcgggcgcag	aagaggccac	gcagaccgtg	gcacaacact	tcgccgacca	106200
ggggcgccgg	acgacccgcg	tgccgggtctc	cgatcgcttc	cactcgccgc	tgatgattcc	106260
gatgtggcg	gagttccgtg	cgggtggcgg	aggactgtcc	tacgccacc	cgtccctccc	106320
cgtcgtctcg	aatctgacgg	gtgggtggc	cacggccgac	gaactgtgct	cggccgagta	106380
ctgggtgcgc	cacgtccgcg	aggcggtccg	cttcgccgac	ggcatcacca	ccctcgaagc	106440
cgaggcggtg	cggaccttcc	tggaactcgg	ccgggacggc	atcctgtccg	cgctggctca	106500
gcagtccctc	gccggcgaag	ccgtcacctg	gcccgtcctg	cgcaaggacc	gcggtgagga	106560
gtccacggcc	ctgacggccc	gagcgcatct	ccacaccgcg	ggactgatcg	aagactggca	106620
ggacttcttc	gtggtgtgg	gtgcggggcg	ggtggagttg	ccgacgatg	cgttcacg	106680
tgggtggttc	tggccggttg	gtcgggttg	tggtggtggt	gatgtgggtg	ctgtggggct	106740
tgggtctcg	gggcatccgt	tggtgggtgc	tgccgtggag	ttggctgcgg	gtgcgggggt	106800
ggtgttgacg	ggtcgtctgt	cgttgctcgtc	gcattggttg	ttggctgatc	atgcggtgat	106860
ggggcggggtg	tttgctcctg	gtacggcggt	gctggagatg	gtgatgcgtg	ctggtgatga	106920
ggtgggggtg	ggtcgtgttg	aggaaactgac	gttggcgggc	ccgttggtgt	tgcttgagcg	106980
tggtgggggtg	cgggttcagg	ttgctgtgga	tgctcctgat	gctgcgggtc	gtcgtggtgt	107040
gggggtgtat	tcgtgtcctg	atggtgtggg	tcaggcggtg	tggtcgcagc	atgctgtcgg	107100
tgtgttggtc	tctggtgcgg	ctgaccaggt	cgggtgggttc	ggtgacggtg	gtgtgtggcc	107160
gccgcagggt	gcggtgtcgg	tggatgctga	gggtgctac	gagctgtttg	cggatgctgg	107220
gttcggttat	ggcccgggtg	tccaggggtt	gcgtgcgggtg	tggcgtcgcg	gcgaggaaact	107280
cttcgcagag	gtcgccctgt	cggacgaggt	tgctgagagc	gctgatacgg	cgaccggttt	107340
cgggttgac	ccggcggtgc	tggatgcctc	gtccatgcc	tcgcttctct	cctcccttga	107400

aggatcaatcc	gccgatgggtg	ggcctgcggtt	gccgttcgcg	tgggaggggtg	tttccctctt	107460
cgctcgggt	gcgacggctt	tgcgcgtgcg	ggtggcgccg	gcgggtgagc	atgcggtgtc	107520
ggtgaccgcg	gtggatccga	ccggtgcgcc	ggtgatttcc	atcgacgcgc	ttcgtaccgc	107580
tcgcctcacc	ctcgatgagg	tcaacgcac	ccacaccag	ctgagcgatg	cgctcttcgg	107640
cgccaatgg	accacgggtc	cgagcacc	ggccgcgcag	caccgcgcg	tagccatcat	107700
cggaaaccgac	cccttcggcc	tcgcagacgg	cctttcggac	gccttgcccc	tggtcgagga	107760
gcgcggtgac	ctgcgcggcg	tcgcagcg	ggagcacc	gtaccggacc	tcgtcctcgt	107820
cccggtagcg	ggcaccgcg	gcacaggcgt	acctgcggac	gccgaaggac	acaccgacgc	107880
cgggacatcc	gacatgctcc	gatccgtgcg	tgaggccacc	gcacaggtag	tggagcagat	107940
ccagcagtg	ttggcggacg	accggttcga	ggcggcgccg	ctggtgttcg	tgacgcgcgg	108000
ggcgggttcc	gtgggtgagg	gcggcatcgc	cgacctggcg	gcctcggccg	tctggggtct	108060
ggtgcggtcg	gcgcagtcgg	agaatccggg	ctgcttcggg	cttctcgacc	tcgacctcga	108120
cctgcgcctt	gactccgacc	ttgccccga	ggtcgacatc	gagcgcgacc	gtgaccgcga	108180
tcgggtcggt	gggaccgtgc	agcccgcgct	cgccgcggcc	ctgcacgcga	ccgccgacga	108240
gccgcagttg	gcactgcgcg	gcgggaccgt	gcaggccgcc	cgactgacc	gaatccccgc	108300
gccgcagacc	gaccgtgcgc	agaccgacc	tgccgagacc	gaccgtccgg	agatcgacac	108360
ccggcggccc	ggcaccggtg	tcatcaccgg	tggtaccggt	ggcctcgggtg	ggttgctcgc	108420
ccggcacctc	gtcgcgcgag	gggggggtacg	gagcctggtg	ctcgccagcc	ggagcgggtct	108480
cgcggccgag	ggagcggaga	agctggtcgc	cgacctcgaa	gcgctcgggtg	ccgtggtggc	108540
cgtgcagacg	tgtgatgtgg	ccgatggcga	tgccggtggcg	gcgttggtcg	ccggcgtgtc	108600
ggacagtagt	ccgctgacgg	cggtcgtcca	cacggccggg	gtgttgagcg	acggagtgat	108660
cggctcgctc	accgaggagc	ggctcgccac	cgtcgcgcg	cccaaggcgg	atgccgcctg	108720
gcactctgcac	gaggcgacc	gcgatctgga	cctggacgcg	ttcgtcgtct	tctcctcct	108780
cgcggcgctc	ctcggtggcg	ccggtcaggc	caactacgcg	gcggcgaaca	cgttctctgga	108840
cgccttgatg	gcgcagcg	gcgcgcgcgg	ggtgcccggg	gtgtcgctgg	cgtgggggtcc	108900
gtgggaccgg	gccggcgcca	tgacggggac	cctgtcggac	gccgaggccg	accgcctcgc	108960
ccgctccggt	gttccgcgga	tctcggcgga	gcagggcctt	gcgctgtacg	acgcggcgac	109020
cgcgggtgag	cggccgctgg	tggtgccggg	gcggctggac	ctcgccgcgc	tccgcgggct	109080
cggtgatgtc	ccggcgctgc	tgcgcggact	ggtccggacg	ccgcgcgcgc	ggaccgcggc	109140
ggccgggtgcg	gcgcgcgcgc	ccgatgtgct	caccgcgcag	ttggccgggc	tcggcggggc	109200
ggagcaggag	gaggtcctgc	tgaggctggt	gcgcggtcag	gccgcgggtg	tgctcgggca	109260
cgcgcagcgg	tcggcgatcg	gtgcggggcg	acagttccag	gagttgggct	tcgactcgct	109320
gaccgcgggtg	gagttccgca	accgactcaa	cgcggccacc	ggactgcggc	tgccggccac	109380
cctgctgttc	gactaccgca	cgcggccgga	cgtcgtcggg	cacctgcgcg	gccggctcgg	109440
caccggggag	gtgtcgggtg	cgggctcggg	gctggcgggc	ctggacaacc	ttgaggcggg	109500
gatcgccggg	ctgtccctcg	acgacgcggg	ggagcaccag	ttggtggccg	gccgggtgga	109560
ggtcctcagg	gcgaagtggg	cggacatgcg	aagcgcggag	ggagctgtgg	acggcgggtgc	109620
ggacgtcgac	atcgaggagg	cgtcggacga	cgacatgttc	gcgctgctgg	acgacgagct	109680
ggggctgaac	tgagccgctc	cgcgatgagca	gttccgcacc	agaagtcca	cagtggcctg	109740
gtccacagtg	acctggtcca	cagtgaagctg	ttccacagcg	acctgttccg	cggcgggttcc	109800
gtagtgagcc	gttccgctta	ggcgggaac	cggggccccg	gtggtcggac	gcttgattcc	109860
ggccaccggg	agcgtcaaac	cggcctcttc	taaagaagg	gaaagaacta	gggaacacc	109920
cacagcccga	tcagtgtgaa	tgaattccgt	ggctcggaaa	accattcgcg	ggccatggag	109980
ttacgggtgtg	atcacgtgtc	cgtacggcag	tgacggcggtg	cacgggttgc	cacggcaggt	110040
ccccggccga	ggcgtctccc	cgtcgtcgc	cggcgggtgcg	tgccgcttgg	ttcgtcccta	110100
ggaggtcccc	ccatgaccac	gtccaccgag	gagagcctgt	gggcccgggtg	cttccatcca	110160
gcaccggccg	cccccgctcg	gctcttctgt	ttccacatg	cgggcgggtc	ggcctccttc	110220
tacttcccgg	tgctggccca	actgtcctcg	ggtgcgcagg	tggtcgccat	ccagtaccgc	110280
gggcgcacag	accggcgcaa	ggaagccggg	gtcagtgacc	tcgcgacctt	ggccgaccag	110340
gtctacgacg	cgtcgcgcgc	cctgctgaag	gagcggccga	gcacgttctt	cgggcacagc	110400
atgggcgcga	cgtcggcctt	cggggtggcc	cggcgcttcg	aggccgacga	cgggtgacctg	110460
gtccggctgt	tcgcctccgg	gcgcggggcc	ccctcccgcg	tgctgaaga	ggcgtgcac	110520
cggcgggtccg	acgacggcat	cgtcagggag	ctgaagctgc	tcgcgggcac	caacaccgcg	110580
ctgctcggcg	acgaggagat	cctgcggatg	atcctgcccg	cgatccgcag	cgactaccag	110640
gccatcgaga	cctaccgctg	cccgcgccgac	gtcaccgtcc	gggcgcgcgt	gaccgtcctc	110700
accggcgacc	gcgaccgaa	gacctccctg	gacgaggccg	aggcgtggcg	cggccacacc	110760
accggggact	tcgacctcaa	ggtgcttccc	ggtgggcaact	tcttcgtcag	ctccgaggcc	110820
ccggcgatca	tcgatctgct	ccgggcgcac	ctgcgcggca	acggctagcg	ggcgcactgc	110880
ggcaggccgg	cgggtgcggg	ctgcgcgcag	gccccgcgc	cgctgagac	ggcaccatgc	110940
caccggcgga	cgcgcgttgc	gtgcgcctgt	ggagcaccgt	ccgcgctgcg	tacgcgggcg	111000

cgcgaggagac	gtgctcgccc	ggccggatgt	cctggaggag	accattcgga	aaacgtcaaa	111060
tgtctgacag	ctgctgctca	tttactccag	cccgcactat	atgatcttgt	ggtggggcat	111120
gtacggagcg	tgccactcgt	tcgctcctgg	tgcattccatc	ccgaccgtcc	ggatttgtgt	111180
gcaccggacg	gccgtgggtcc	tggtacgcgt	cggcgcgagg	tcggctcctg	ttctcaccce	111240
acgtggagga	acgctgatg	cggagcagt	cgggttcttc	cggcttactg	accacgctgg	111300
tgggacgcga	cgacgaactg	cgcaccctcg	cccggcacgc	cgcggccgcc	cgcgacgggc	111360
gggcccggct	ggtcctgctc	cacgggcccc	cggcatggg	caagaccagt	ctgctgcggt	111420
ccttcacggc	gagcgatgtc	tgccgcggca	tgacgggtgt	gtacggcacc	tgccgagaga	111480
ccgtcgccgg	cgcgggttac	ggcgggtgtc	gcgaactcct	cggcgggctc	ggcctgagcg	111540
gcggcgacgc	ccggcgctcg	ccccctctgg	agggcctggc	ggcccgcgcg	ctgcccgcac	111600
tcaccgcaga	ccccgccggg	cccgacgcgc	ccacgggtgc	ctaccgggtg	ctgcacggcc	111660
tgtactggct	cgcgcgccgc	ctcatggccc	aacggccgct	ggtcctcgtc	ctcgacgacg	111720
tccactgggtg	cgacgaacgc	tccctggcct	ggatcgactt	cctgctgcga	cgcgccgagg	111780
acctgccgct	gctggctcgtc	ctggcctggc	gcagcgaggg	cgaaccgggtc	gcgcccgcgg	111840
tgctcgccga	catcgccgcc	cagcgccgcc	ccaccgtgct	cggcctgcac	ccgctcgggc	111900
ccgacgacat	aggcgaaatg	gtgcgtcgcg	tcttccggac	cacggccgca	ccgtcgttcg	111960
tcagccgggt	cgcgcgcgtg	tccggcggca	atccgctggc	cctcgcccgc	ctcctcgacg	112020
aactccgcgc	cgagggcgctc	cggccggacg	ccgcggggga	gcgcggggcc	gccgaggtcg	112080
gcagtcacgt	cctcgcccgc	tcgggtgcgt	gcctgttggg	gcgcggggcc	ccctgggtgc	112140
gcggcgtggc	ccgtgccatc	gccgtactcg	gcccggagtg	caccgagttg	ctggcgggcg	112200
tcgcccggct	cccggccgcg	accgtcgacg	aggccctggt	ggtgctgcgc	aggccgggca	112260
tccctggccgc	cgaccgcgtg	gacttcgtcc	atgacgtcgt	ccgctccgcg	gtgctcgacg	112320
acgtcgcccc	gcccaccctg	gccgaactgc	gcaccaacgc	cgcgtgttg	ctgagcgacg	112380
cgggcgcgcc	ctccgaggag	ctcgccggcc	agctcatgct	gctgccgggtg	ctcgaccagc	112440
cgtggatggc	cgcggtgctg	cgcgacgcgc	cgcgccaggc	ggagagccgc	ggcgccccgg	112500
aggccggtgt	gcgtgcctc	taccgggtgt	tggaggtgga	gccggacaac	gttgccgtcc	112560
gcattccagat	ggcccgcgcg	ctcgccgaga	tcaaccgcgc	cgaggcgatg	cgcctcctca	112620
aggaagcgct	ctccctcgcc	ggggacgtcc	gcaccgcgcg	ccaggtcgcc	gtccagtacg	112680
gcttcacctg	cctcgccgtg	cagggaatccc	cgtccggggg	gcggatgctg	gaggacgcgc	112740
tcgcccagct	gaccgccgaa	ctgggccccg	aaccagggcc	cgtggaccgg	gagttgcgga	112800
cctcgtgga	gtcgtgctg	ctcatcgctg	gggcccagca	gaaggtagcg	atcgggtgcg	112860
tccgcgaccg	ggcgccccgg	ctcaccatgc	cgcgccggca	cacaccggcc	cagcggcaga	112920
tgctggccat	gaccaccgtg	ctgaccgcga	tggacggccg	ggacgcccgg	tcggccgtcg	112980
accaggcccc	ccgcgcctcg	cgcgcacccg	gcgtcgagct	ggaaccctgg	tcgctgctgt	113040
ccgcctcctt	cgcctctctc	ctggccgacg	aggtcgccga	cgcgcagtag	gcactggacc	113100
tcattgctcca	gtacggccag	gacaacgcgc	cgggtgtggac	gtacgtcctg	gcgctctcca	113160
cgcgcgccct	gctccaccac	gggggtggcg	ccttccccga	ggccctggcc	gacgcccaga	113220
ccgcgctcga	gattcctcgg	gaggagcgct	gggcccagcg	cgcgctgctg	ccccgtgtcg	113280
cgttgccac	cgcctcgtc	gaccgcggcg	agcccagagc	cgcggaacac	gtcctcgacg	113340
gcattaccgc	ccccgcctc	gaacgcttcg	tcattcgaata	ccactggtag	ctccaggccc	113400
gcgctacgc	ccgctgggtc	cgcggggatt	tccaaggagc	cctggacctc	cttctcgctt	113460
gcggctcggtc	cctggaggag	tcgcgcttca	gcaaccgcgc	gttcgtgcgc	tggtggggcg	113520
atggcgcggt	gctcctggcg	accctggacc	gccacgacca	ggcgcgcgaa	ctcgccgcat	113580
acggaagcga	gttgggggga	cgttggggga	cggcgcgcg	cctcggaactg	gccttcattg	113640
cccaggcggt	ggccgcaccc	ggccgcggcg	gcattcgatca	cctcaccgag	gcggtctcgc	113700
tgctcgcgga	ctccccggcc	cgggcccattg	aggcccgggc	cgaacttctt	ctcgacacg	113760
cccacctgaa	gcgcgacgac	ctgcggggcg	cccgggaaca	cctgcgcgcc	gccgcccacc	113820
tcgcccagcg	ctgcggcgcc	gtgaagctcg	gcgtcgacgc	cagaaaaactg	ctggtcaccg	113880
cgggtggtcg	ggtacgcagg	atgaccgcct	ccccactcga	catgctgacc	gggatggaac	113940
gcacgggtggc	ggacctcgcg	gtgaccggcg	cgagcaaccg	ggccatcgcg	gaagccctct	114000
tcgtgactgt	aaggaccatc	gaaacccatc	tcacgagcgt	ctaccggaag	ctcggggtcg	114060
gcgggcgtgc	ggagctgtcc	gccgtcctgg	agaccaggac	cgcacacctc	ggtcggcagc	114120
cgcggccctg	ggtctcccag	gcacgcggac	gcgcttgaga	cgaacaggac	gagaggtagc	114180
ggtgccccgg	agcaaggcgc	ggaaccaacc	gaccacctgc	acgccgcagt	gcgcacccga	114240
cgcgccacggc	gacccccacca	tgctcctcga	atgcggcccg	gaacagcggc	tcattcggcga	114300
cctcctgcac	cgcctcgggc	agggggcgcc	atcggtgctc	agcctgaccg	gccggcccgg	114360
gcacgcccag	aacgccctgg	tccgctgggg	cgcgtgcgcg	gccaggcacg	acgggtcgcg	114420
cgtcctgcga	gcccaggcga	cgcgcgcgga	acgggaactc	cgtacggcg	ccgttctcca	114480
actgctggcc	gtcctcgacg	gcccgcacgg	cagcaccctg	gacgccgcga	tccgccacga	114540
cgttcccccg	ccactgcccg	tgcccggcat	caggaggtg	ctgcggcgca	ccggcacggc	114600



accacacctg	gtcgtggtcg	aagacgtcca	gtgggtggac	ccggcctcgc	tgaegtgggt	114660
gcagatcctg	ctgcgccacc	tggggccgga	caccccgctc	gccgtcctgg	ccagcagctg	114720
cggtgacacc	acggccttcg	acaccgaccc	gaaggccccc	gccgtcccgg	ggccgcggga	114780
caccgtgccc	gtcgcgcgct	tctgtgtgcc	cgcgctcacc	gaccgcgggg	tcgccgccac	114840
cgctccgcgc	gtctgcggca	ccccggcgga	cgaggagttc	atcgccgcgc	tcacctccgc	114900
caccgcgggc	aaccccgcca	tcttgcggga	cgccttgccg	gccttcgtcg	accacggcct	114960
ccccgcggac	gccgaccacc	tcccggagct	gcacgccttc	accgtggcg	tcgtcggcga	115020
ccacaccgtg	cgcgccttgg	acggcctgcc	cgcgaagtc	aacgccgtcc	tgcgggcct	115080
ggcgtctgc	ggcgacctgc	tcgacttcca	ccgagtccgg	gccctcgccg	gcgcgcactc	115140
gctgtccgag	gaccggatcc	gcacctgct	ggcgagcgtc	ggcctgaccg	tgtccgtcgg	115200
cgacaagggtg	cacatccgct	tccccgcctc	caaggcacgc	gtcatcgagg	acatgcccgc	115260
cgcggagcgc	gccgatctgt	acgtccgcgc	ggccgaactc	accacagtt	gcggcgtaa	115320
cgacgaggac	gtcgcctatc	tgtgtgtcgc	ctcgtcgccg	ctcggcgcac	cctgggtcgt	115380
gcccctgctc	cgcgcgggat	tcgcgcgcgc	gctgcgcggg	gaggaccacc	accgggcctg	115440
tgcctgcctc	tcccgcgcgc	tgcaggaacc	cctcgacccc	cggaacgca	gcctgtgac	115500
ggtggaactg	gccgcggccg	aagccgtcgc	ccggccggag	gcgggggatc	gacgcctggg	115560
ggaactcgtc	cgcagcaccg	tcgcggacac	cgacccacag	tcgtccgggtg	aggggggtggg	115620
ggtccgcgc	atcgacctgg	ggttcgcccc	gggcaacagc	gaatgggtcc	gccgcaccgc	115680
gggcgaggcc	ctgcctgacg	ccgggcccgg	cgaccgggag	gaactgggtc	cgtgtttctg	115740
ggtggccgc	gtgcgggacg	acgacgcgc	gatgatcccc	gtggtgcccc	ggttgccccga	115800
ccggccgggtg	ccgcggcccc	aggccggcgc	ccgtgcctgg	cagctggcca	cggcggggga	115860
ggacgcggac	aaggccagga	agctcgcccc	gatcgccctc	accggcgggg	tgaacgagag	115920
cctgatgatg	ccgaaactgg	cggcctgcgc	cgcgtgttcc	gccaccgacg	acaacgacga	115980
ggcgtgcaac	ggcctggaca	ccatgtctac	cgcgcggccg	agtgccacc	tgcgcagcat	116040
ggccgcccgc	attttcaacc	tacgggcccg	gatacacctg	tgcgcggccc	ggctgagggc	116100
gcgcgaacgc	gatctggaca	gcgccgagcg	gcacctgccc	ccgacgagtt	ggcacccccg	116160
tgcgtctccc	aacctgatcg	ccaccgcgat	cctcgtcagc	atggagacgg	gccgcccgga	116220
ccgggcccgc	cgactcgccg	aggccccggt	ccccgcgggc	ggcgaggagg	gtgtgtgggtg	116280
gcccgccttg	ctgctcgccc	gcgcccgtgt	ggccgcggac	gacggtgact	gggaggaggc	116340
cctgcggctg	tcgcgggagt	gcgggcgctg	gctttttcgc	cggcactggg	ccaaccgggc	116400
catgtctcagt	tggcgccccg	tggccgcgca	ggcgtgtctg	aagctcggcg	acgtgacgga	116460
ggcgcgcggg	ctgcgggacg	aggagctgtt	cttcgcggac	cgtgggggca	ccgcgagcgc	116520
ccgcgggatc	gcccgcctga	cgacgcggcg	actcttcgac	gacgacggcg	accgggccgt	116580
ccggcggatc	cgcgaggccg	ccgccctgct	ccgcgactcg	cccgcgccgc	tggcctacct	116640
gtggagccgg	ctgagccagg	ccggtgccga	gacggcccac	ggcgacaccg	ccgcggccgc	116700
acgtctcctg	caggcggtcg	cccggatgac	cgccgcccac	cccgccagcc	gcctcgccac	116760
cgccgccccg	accctgaccg	tcccgtccgt	tccggtcgcc	accgcgccgc	ccaccgccgt	116820
cgtcccaccc	ggatggcgcg	acctgtccga	ggcggagaag	gacaccgtgc	tgtcgcggcg	116880
ccgcggccac	ggcaaccgcc	agatcgccga	acaactcgcc	gtcagcaggc	gcaccgtgga	116940
gctccggctg	agcaacgcct	accgcaagct	gaggatcggc	ggacgcaagg	agctgtacct	117000
gctcctggag	gcgttggaag	gaccggtcgc	ggatgcttct	tgagcgggag	aacgaactgg	117060
cccggatccg	ggcgccttg	gacgcgcggg	aagcgggcca	ctcctcgctc	ctcctgatca	117120
acggtccctt	cggcagcggg	cgttcggcgc	tgtgcgcggg	gataccggag	ctggccggcg	117180
acggcacccg	cgtcctgcgg	gccagcgccg	cctggcgggg	acgcgacttc	cccttcggga	117240
tcgcccggca	actcttcgac	cacctgctgt	ccggggcggg	cggcgagggg	ccggccgaac	117300
gcaccgcggg	ggcagagcac	ttcagccgac	tgatggacac	cggcgaccgc	cctaccggga	117360
ccggccccgc	cctggaggtc	tcccaggcag	tgtccagggg	cggccaggcg	ctgctcgccg	117420
acgcgtccgc	ggagcggcgc	ctgctgatcc	tggctgacga	cctccagtgg	gccgacggcc	117480
cgtccctgcg	ctggctggcc	cacctacccc	ggcggtgca	cggcctgcgg	gcgctgctgg	117540
tgtgcacgct	ggccgacggc	gaccaccggg	gcaggtaccc	cctggctccg	gaggtcgccg	117600
gcgcgcgcga	caccgtcctg	cgcctggcgc	cgtgttcccg	ggacgccacc	cgcgtcctgc	117660
tcgccggggc	ccagggccgg	ccgcgcgagg	acgcactggt	gcgcgcctgt	tacgaggcgt	117720
ccaggggcaa	cccgtgttcc	ctgaccgcct	tccggagcgc	tctgcgcgcc	accggaaggc	117780
cgcgcggcgg	cgaccacttc	ggcgccgtcc	gggagctgag	cccgacgggtg	ctgcgcgatc	117840
ggctcgcggg	ccatctgcgg	atccagccgc	agccgggtgcg	cgaggctcgcg	gtggcggtgg	117900
ccgcgctggg	cgaccacagc	gatccgggtg	tgtcgcacca	gctcgcgggg	gtcgatgaaa	117960
tcggtttcgc	cgggtgccgc	cgcgcgctgg	tggacgcggg	cctgttgggc	cggggacggg	118020
acgtccgctt	cgtccacggc	gtcgtccgcg	atgcgggtgga	ctccctgctc	accctcgacg	118080
agcgggaacg	ctcgcacgac	gacgcgcgcg	atctgctgta	ccgctgcggg	cggccggccg	118140
agcaggtcgc	cggccatctg	ctggccgtgg	tccaccgggg	ccggccctgg	tcggaggcgg	118200

tectgcgctc	cgcgccccac	aacgcgctgc	gcgcgcggcg	gcccgcgcag	gcggccccggt	118260
acctgcgccc	cgccctgctg	caccaccgca	cccaggacgg	ctgcgcgcgc	cgcatacctgg	118320
tcgatctggc	caccgcgcgag	cgcgccctcg	accccgatgc	ctgtgtacgc	cacgtcagcc	118380
aggcggtcgc	gctgctggac	acctcgcggg	atcgggccgc	cgccgtggtg	cgcataccgc	118440
cgccccctgt	cgccgcccc	agcccgtccg	ccgtcgagtt	ggtgcggcag	gccgcgcgcg	118500
ggctcgacga	accggggcag	cgggacgagg	agggagccga	cgaactcgcc	ctgcgcctgg	118560
aggcgtggct	gcggcactcc	ggccacgaga	accccgtcga	gctggcgctc	tcgggtggcgc	118620
ggctgcggcg	catgggggca	cggccgcgcg	tggacagcgt	cgccgaacgc	gagttggctcg	118680
ccgtgctggt	gagtgcgggc	gcgctcagcg	gccggctcag	cgcgcgcggag	atcgccgaca	118740
ccggcaaccg	catcctcgaa	cgtgagccgg	ccaccgcgcg	ccatgcccac	accccgctgc	118800
cgctgggtgat	gctctcgctg	ttcgctcgccg	agtccgtgca	gggcgtggcc	tcctggctgg	118860
ccagcgaaca	gcacaccggg	cgccgggtacg	cgaccggcgc	ggacgatgtg	ctgctgaccg	118920
ccgagcgggg	cttcgtcctg	gtgacgcagg	gccgcccggc	cgcgcgcgcg	gagcacgtcg	118980
aacgcgccct	ggtcatggac	gccggcgact	ggtcggaacc	cgcgcgtcatg	atgttcgcgc	119040
cggtcgcctt	cgaactgcgc	gaccgcgcct	tgagcgaacg	catcctggaa	cggatccgcg	119100
accgcgcggc	ggccggactg	gcgctcaccg	ccaccgggtca	gatgctccag	gccgcgcgtcg	119160
acgtgcactt	cggccggggg	cgggacgccc	tggacacgct	gctggcctgt	ggccgacgcc	119220
tggagaccgt	gggatggcgc	aattccgcac	tgctgcctcg	gcgtccgtat	gcaatcgggc	119280
tgcaccagcg	gctcggcgag	accgatgccg	cgctgcaact	cgcgcgaggac	gagctgaggt	119340
gggccccgga	gtgggtgctg	acgacgaacc	tgggcccggc	cctgcgcctg	aagggtggc	119400
tgttgacgga	cgagggactg	gatctgctgc	gtgagagcgt	cgagatcctc	cgcgcttcgt	119460
cctacgcgac	ggaactggcg	cgcacccctcg	tcgtcctcgg	gcgtcggctg	ccgggtgggc	119520
cggagggccga	ggcgggtgtg	cgggaggccg	cggggatcgc	cgcgcctgt	ggtgttcctt	119580
ggctcgccga	acgggcccga	ctcggctcgg	gcagtgccat	cgtgccgcgc	gtcgccaccc	119640
tgaccccgag	cgagcggcga	gtggcgctgc	tggtagccg	gggcctgacc	aaccaggcca	119700
tcgcgaccga	actcgggtgtg	agtcgccggg	cggtaggaga	gcacctcacc	agcgcctatc	119760
gcaagctcgg	cgtctccggc	cgcgcgcagt	tggtagaatg	gctcccgggt	cgttgacgcc	119820
cgcggccgctc	ccgttcggga	cgatcacaa	tcctgtgtgc	tgtagccct	gtgggtgcgg	119880
cgctcgtgtc	gtacgcccgt	ccggtcatga	cagcaatcct	caacaaaagc	tcaaacggaa	119940
gcgggggcag	gtgcatccac	cccatcacca	atgtctcgta	tggtactttt	ggcccaccgt	120000
agttgggact	ctgcggaccg	tacctcgta	agtacggacc	acggatttag	gggtgcgtag	120060
atgtgtcgcg	ttgacgtgcc	gagttgtggc	taactacgct	ccgtttcacg	gaaatcgaaa	120120
ctcttcggcg	tgaatcgcg	cgaacttcga	tcgaattctg	gattccgttt	ctcagaaggc	120180
taaaagacga	cgggggttat	ccgaaaggat	ggcgggtctg	tgactatac	tcacctact	120240
gactcgaatc	accgtacgac	cggcgggggtg	atatccgctc	agaccgcacc	ggccggggag	120300
tcctgcggcc	ccgggctgat	ggcgctccctc	gaccgtgacc	tcaccatcaa	gcacgccaac	120360
caggagttcc	gccgcgcgtt	cgacgattcc	gccggtgacg	tctgcgggtcg	cagcttcggg	120420
gacctgatgc	acccgagcgt	gcagcagccg	ctgatgcgcc	agttctcccg	gctgatcgag	120480
ggcaagcggc	accgcttcgc	ctcgcacgtg	gtcgccgtgg	gcgcccagga	cgcggccttc	120540
gcgggcaccc	tcaccgcctc	cgcggtcacc	ggcaagaccc	cggacatcgc	cgggatcctg	120600
gtgctcatgg	actcgtccgg	cgcgcgcgac	gccgcgatg	ccggcgctcg	gaccagccag	120660
aagaagtcc	tgaccgagat	cgatgcgcgc	atcctcgaag	gcacgcgggc	cgggctgtcg	120720
accattccgc	tcgcctcgcg	gctctacctc	agccgcgagg	gcgtggagta	ccacgtcacc	120780
gggctgctgc	ggaagtgtcg	ggtgcccac	cgcgcgcgcg	tcgtctcgcg	cgcctactcc	120840
atgggcatcc	tgaacgtcgg	cacctggccc	ccgaaggctg	tggacgactt	catcaagtga	120900
cgcgcgcgcg	gcctgtcccg	caccggccgg	ccaccgtcgg	accgctgtcc	cgcacgcccg	120960
tgctgcgcag	gccgcctct	cacggggcga	tgccggcgcg	cacgggcgtt	tcgtgccctg	121020
gtcgatgttt	atccattctc	acgttacgaa	cccgaatgc	tcagcgcgcg	gattccggcc	121080
acttttcggt	ccgtcgtaac	caaactgccc	gtcatatcgt	ttccactgac	cgtttctgtg	121140
gtccgcgggtg	tgccacctgg	gccccactcc	gacgcgcccg	cgacgggtcg	cgagcggcct	121200
ccgaatactc	acggtaaccc	ccgttcctcc	tggtagagtt	ccccgcgcgt	cacattcctc	121260
cgtgcctcgt	gcggcggtgg	tcgcccgtt	tcgccgcgcg	cgcgacagcc	attgtttacaa	121320
cggcacgaca	gagcgggggt	ggtcgaccgc	tcgcgggtgc	ggcgagtcga	agcggaaacct	121380
ctgctcgatg	ttcggtccac	cccgaactct	gtgttcttgg	tgaccgtttt	caccaatgcc	121440
ccgcccgttc	ggtggcgaat	ccgcgcgcgc	ttcctggcca	caacggatcg	cttcggccgc	121500
ccaccacggc	cggcgagcc	gaggcttgtt	ggaatcacc	ccaccagtc	ggggcgccgt	121560
tgtggcccgg	acacgcagga	ggaacaccgt	ggatgctgag	ggccgcgcga	gggacatgct	121620
ggagctgatc	cggcgagcg	gatccgcgga	tgctcgctgc	ctcgccgagg	agttcgccgt	121680
cagcaaggaa	acggtccgc	gggacctgaa	cgctcctggag	ggccatgggc	tgatacggcg	121740
gcggcacggg	ggcgccctatc	ccatggtgcg	gccgggctcc	gaggccgtct	tcgtctcccg	121800



gaccgcacag	ccgatccccg	aggagtcccc	gatcgccacc	gcggccgcgc	aactcctcag	121860
cgaggccgag	acggtcttca	tcgacgaggg	cttcaccccc	caactcatcg	ccgacgcctt	121920
gccgcgcgac	cggccgctga	ccatcgtcac	cgcctcgctc	ccggtggtca	gcgccttcgc	121980
gacgagccca	caggccaacg	tgctgctcct	gggcgggccg	gtccgcccgg	gcacgacggc	122040
caccgtcgac	cactgggccc	tccacatgct	gtccggcttc	gtcatcgacc	tggccttcct	122100
cggcgccggg	gggatctcgc	gcaggtacgg	cctgaccacc	cccgaacccg	cggtcgcccga	122160
ggtcaaggcc	caggccatcc	gcgtcgcgcg	ccgcccggtc	ctcgccgggg	tgacacacaa	122220
gttcggcacg	gcgagcttct	gccgggttcg	agaggtgggc	gacctggaga	cgatcgtcac	122280
cggcgccggc	ctgcccgtcg	ccgaggccca	ccgtaccac	ctcatggggc	ccaaggtttt	122340
acgggtgtga	cgcgcgcggg	gcgtcccggc	gcccgcacac	ggcctggcgc	gcgcgggtgc	122400
gggtcagccg	cgtcgcgcgt	ccgtcggttc	ccgccagagt	gcgagcggtg	tcacccccac	122460
ccggtagccc	tcctcgtcga	ggagatggcc	gccgtgcttc	atgtgccaga	gggcgacccg	122520
gcgttcgggtg	acggtcaggt	ccgcggcccc	ttcggcgagt	cggacctcga	agcgctgggc	122580
gtcgtcgagg	gagcgcaccc	aggccaccag	gtgcagggtg	tgccggctga	tgacgtcgc	122640
gcacaggcgg	acctcgcgca	tcgccgtgac	ccggcgggtc	acctcgcgca	gcctcgcggc	122700
cgggacctgc	ccccagaagc	tcaccgtgac	cggccactcg	gacagcgggc	gggcgacctc	122760
gcaccggggc	tgcaacatgt	ccgcggcgaa	gagccgttgg	acgcgacggc	gcacggtgtc	122820
ggggccggcg	ccgcactgct	cggccagcgc	gcggtaggtg	gcccggccgt	ccacggacag	122880
cgcggtgacc	agctgtttgt	cgagctcgtc	gaggacgaag	gcgggggtgt	ccgtgcgggtg	122940
tctggacgcg	tcggcgccga	gccggggcgac	ctggtgccgg	ccgagggccc	gcaggcgcca	123000
ccggtgccc	tcggtgtgca	ccgggcccgc	caggtgcgtc	cgggcggccc	ggacgccgtc	123060
caacgcggcg	aggtcgtggg	tcacccaacg	ggagagcatg	gccggatcgc	gggccatgac	123120
gttgagttgg	aggtcacggt	ccccgggtgac	gtgggagagc	gccaccacgt	gcggcgccgc	123180
ggccaacgcc	cgcgcgacgt	cgagcagtcg	gccgggagcg	cagtcgatct	cgacgaacgc	123240
caggcacccc	tggcccgcag	cggccagcac	cggggccgga	tggcagctga	tccaggcggc	123300
gccggtctcg	accagccggt	tccagcgctt	ggccaccgtc	acggcgctca	gtccgaggac	123360
ggagccgata	cgggtccagc	tggcccgcgg	cgtgatctgc	agcgcggtga	ccagggcctg	123420
atccacatgg	tccagggacc	gggggggtctg	cccgggaatcc	tgcgccacgg	gccacctcct	123480
tgcggtgttc	cggcggtatc	gggcccgcgg	tcgggtcaac	cttcagcctg	gactcgggta	123540
cggccggacc	gtaccaggca	acccccggag	caacaggagt			123580

//